J Comp Physiol A (2015) 201:563–584
DOI 10.1007/s00359-015-1002-y

Three-dimensional models of natural environments and the mapping of navigational information

Wolfgang Stürzl[1], Iris Grixa[1], Elmar Mair[1,] Ajay Narendra[2,3], Jochen Zeil[2]

[1]German Aerospace Center (DLR), Institute of Robotics and Mechatronics, Münchner Str. 20 D-82234 Wessling, Germany

[2]Research School of Biology, The Australian National University, Bld. 46, Biology Place, Canberra ACT 0200, Australia

Present address: Department of Biological Sciences, Macquarie University, Sydney NSW 2109, Australia

*Corresponding author: wolfgang.stuerzl@dlr.de

Keywords: Visual navigation; insect homing; 3D models; natural navigation environments

Abstract

Much evidence has accumulated in recent years, demonstrating that the degree to which navigating insects rely on path integration or landmark guidance when displaced depends on the navigational information content of their specific habitat. There is thus a need to quantify this information content. Here we present one way of achieving this by constructing 3D models of natural environments using a laser scanner and purely camera-based methods that allow us to render panoramic views at any location. We provide (1) ground-truthing of such reconstructed views against panoramic images recorded at the same locations; (2) evidence of their potential to map the navigational information content of natural habitats; (3) methods to register these models with GPS or with stereo-camera recordings and (4) examples of their use in reconstructing the visual information available to walking and flying insects. We discuss the current limitations of 3D modelling, including the lack of spectral and polarisation information, but also the opportunities such models offer to map the navigational information content of natural habitats and to test visual navigation algorithms under 'real-life' conditions.

Introduction

In 1996, Rüdiger Wehner, Barbara Michel and Per Antonsen (Wehner et al. 1996) presented the stunning result of an experiment in which two zero-vector ants (*Cataglyphis fortis*) - that is ants that were captured at the nest after returning from a foraging excursion and therefore had their home vector zeroed – were released back at the feeder position they had been visiting about 30 m away from the nest (Fig. 1). After a brief search around the feeder site, both ants practically retraced their steps back to the nest, through a complex terrain of tussock grasses, indicating that they had previously memorized their complete homing paths. Since then it has become clear that ants of different species inhabiting landmark-rich environments memorize multiple routes (e.g. *Melophorus bagoti*: Kohler and Wehner 2005, Sommer et al. 2008) and in addition acquire these memories very rapidly (*Cataglyphis velox*: Mangan and Webb 2012). Moreover, such route memories can in principle be based on very coarse scene representation or classification that still would allow insects to recognize familiar scenes and to determine the appropriate heading direction by scanning at each segment of extended routes (Baddeley et al. 2011, 2012, Möller 2012).

There is also now solid evidence that insects are guided by different navigational mechanisms depending in each specific case on the availability and reliability of navigational cues in their habitat (e.g. Narendra 2007a,b; Wehner 2008; Wajnberg et al. 2010; Buehlmann et al. 2011; Sandoval et al.

2012; Cheng et al. 2012; Collett et al. 2013a; Legge et al. 2014; Wystrach et al. 2014b). In landmark-poor environments, such as salt pan deserts, ants rely on a celestial compass for guidance (e.g. Wehner 1997, Wehner and Müller 2006, Wehner and Labhart 2006), but also on wind direction (*C. fortis*: Wolf and Wehner 2000, *M. bagoti*: Wystrach and Schwarz 2013) and on olfactory landmarks (Steck et al. 2011; Buehlmann et al. 2012). Environments with three-dimensional structure provide animals in addition with terrestrial visual compass cues (Zeil et al. 2003, Graham and Cheng 2009a,b) and most importantly with robust cues to location in space (Zeil et al. 2003, reviewed in Zeil 2012, Collett et al. 2013a), to the extent that in visually structured habitats and over a certain range, insects rely mainly on visual landmark guidance (e.g. von Frisch and Lindauer 1954; Narendra 2007a,b; Narendra et al. 2013a,b).

The extent to which a given natural habitat offers robust cues for visual navigation can be quantified by comparing reference images close to a goal and/or along a route to views that are seen at other locations in the environment (Zeil et al. 2003; Philippides et al. 2011). Insects appear to memorize the scene close to their nest or a food source during learning walks (Nicholson et al. 1999, Müller and Wehner 2010, Jayatilaka et al. 2013b, Dewar et al. 2014) and learning flights (e.g. Zeil et al. 1996, Philippides et al. 2013, Collett et al. 2013b). It is not clear at this stage, whether route views are learnt continuously or depending on how much views change. The navigational information content of panoramic images is two-fold (see Zeil et al. 2003, Stürzl and Zeil 2007, Zeil 2012, Collett et al. 2013a): views change smoothly with distance from a reference location, a fact that is described by a translational image difference function (transIDF) and also through changes in orientation away from the reference orientation, described by a rotational image difference function (rotIDF). Assuming that insects have acquired reference images close to goals or along routes, the navigational information they have available when, for instance, experimentally displaced can be quantified by comparing such reference images with the views experienced at new locations. So far, this has been done by recording panoramic images, unwarping them to rectangular panoramas and by calculating the global root mean square pixel differences between different locations and orientations in experimental arenas (Cheung et al. 2008, Stürzl et al. 2008, Mangan and Webb 2009, Wystrach and Beugnon 2009, Wystrach et al. 2011a) or outdoors (Zeil et al. 2003, Stürzl and Zeil 2007, Wystrach et al. 2011b, 2012, 2014a; Wystrach and Graham 2012; Narendra et al. 2013a,b; Schultheiss et al. 2013, Zeil et al. 2014).

However, in practise it is very cumbersome or even impossible using panoramic cameras to map in different environments the range over which panoramic reference images provide navigational information. This is especially the case when one considers the views experienced by flying insects.

Here we report on developing a tool-kit for constructing 3D models of natural navigation environments with the aim of quantifying navigational information, of mapping the range over which visual guidance is provided in specific habitats by panoramic views, of reconstructing the views experienced by navigating insects and of providing complex, natural benchmark environments for testing homing algorithms. Basten and Mallot (2010) presented an earlier, indirect attempt at building 3D models of an ant environment, by using a published map of an experimental site to construct a virtual model of the area. Similarly, Mangan (2011) constructed a rudimentary model of an ant habitat by using information from a map and from panoramic images (now available at www.insectvision.org).Some procedures and results of our own methods have been described in preliminary form in Mair et al. 2013 and Stürzl et al. 2013 and are publicly available at www.insectvision.org.

Materials and Methods

Experimental sites: We created 3D models of two field sites in Canberra, Australia. One is a small urban park (35°15'05.59"S, 149°09'33.18"E) where we conduct work on the navigational abilities of the jack jumper ant *Myrmecia croslandi* (Narendra et al. 2013a, Zeil et al. 2014). The other site is a nest aggregation of ground-nesting wasps within Mt Majura Nature Reserve, Canberra, Australia (35°14'36.98"S, 149°10'10.56"E) where we study the relationship between learning flights and the homing abilities of these insects.

3D modelling and acquisition of panoramic images: We used two different approaches to create 3D models of the experimental sites which will be outlined in the following sections. One approach was using a laser scanner, which directly measures the distance of objects reflecting the laser within a certain radius around the scanner and outputs, in combination with a colour camera, a coloured 3D point cloud. The second approach was using digital cameras to acquire overlapping images of a certain area, from which by means of a method called Structure from Motion (explained in more detail below), the 3D structure of the scene can be calculated. The output is also a coloured 3D point cloud.

For the first approach, we used a Laser scanner/colour camera combination (Z+F IMAGER® 5006i, with an attached motorized colour camera (Z+F M-Cam), Zoller+Fröhlich GmbH, Wangen, Germany) to scan these two sites from multiple locations in order to minimize occlusions. The system sits on a motorized, levelled platform (Fig. 2a). Scans run automatically according to preset programs with the laser scan followed by images taken by the integrated colour camera along three elevation slices.

Typical acquisition time is about 10 minutes for a full scan. The system is robust and easy-to-use in the field generating high resolution data (angular resolution is up to 20000 points/360^{o}) with a 360^{o} horizontal and 310^{o} vertical field of view (see Fig 2b for an example scan), for a range of distances from 0.5 m to 80 m. To combine several scans into a common coordinate system, but also to estimate the transformations needed for mapping camera images onto individual scans, corresponding 3D points have to be identified in different scans so that rotations and translations for each scan with respect to the reference frame can be estimated by means of non-linear regression or other methods. We pinned markers printed on A4 paper to trees and manually identified corresponding markers in each scan. Zoller+Fröhlich's software (Z+F LaserControl) allowed us to create 3D point clouds from these scans that could be registered with RGB colour data from the M-Cam. We also used a UV camera at the same nodal point position as the scanner, to map UV information into the 3D point clouds generated by the laser scanner. We used custom-written software to reconstruct from these coloured point clouds panoramic views by remapping six 100x100 pixel rendered views to 360x180 pixel panoramic images (1^{o}/pixel resolution, see Fig 12d) within the range of the model at defined positions and orientations. The differential GPS coordinates of four landscape features that were easily identifiable in the laser scans were used for aligning the 3D model with the GPS reference system that we employ to track ant paths and to locate nest and release sites. To reconstruct views from the cockpit of flying insects, we used markers in the ground to register high-speed stereo camera footage with 3D models. The 3D coordinates of markers as defined by the stereo camera coordinate system were mapped into the coordinate system of the 3D model, which contained the same ground markers. The 3D flight path coordinates were thus equally aligned with the model coordinate system.

Camera-based 3D reconstruction from RGB and UV images: Laser scanners become cumbersome or impossible to use when a detailed reconstruction of complex ground structures is required. For such fine-scale modelling around the nest areas of ground-nesting wasps, we reconstructed an area of approximately 2 square metres around the nests using purely camera-based methods that only require image series from hand-held cameras and rely on Structure from Motion algorithms (Hartley and Zisserman 2003). The images were taken with an off-the-shelf digital cameras (Canon IXUS 220HS, Panasonic DMC-FX200) with 'focus lock' enabled and a UV camera (CM-140GE-UV, JAI, Yokohama, Japan) to account for the insects' ability to sense light in the blue, green and UV spectral regions (e.g. ants: Labhart 1986; Ogawa et al. 2015; insects: Briscoe and Chittka 2001). In the next section, we describe our workflow from the camera images to the full 3D blue-green-UV model using either open source or freely available software or Pix4DMapper by Pix4D (Lausanne, Switzerland).

Given the rapid development of 3D computer vision in the last 20 years, image based 3D reconstruction methods are being applied in various fields such as cultural heritage preservation, architectural modelling and recently also in studies of animal locomotion (e.g. Pollefeys 2004, Snavely et al. 2006, Wohlfeil et al. 2013, Sellers and Hirasaki 2014). Such applications have also benefitted from free tools that implement Structure from Motion (SfM) techniques, such as *Bundler*, *123D catch*, *VisualSFM* and others. A big advantage is the fact that no special equipment apart from a camera with a lens that fits the projection model of the software is needed for a reconstruction. Any images taken from any camera or set of cameras to which the perspective camera model can be applied will suffice.

Structure from Motion is the process of recovering the optical geometry of a set of cameras and their positions and orientations for a given number of images taken from multiple viewpoints, while simultaneously reconstructing the 3D geometry of the scene (Hartley and Zisserman 2003). No *a priori* knowledge is required of camera positions or the 3D location of reference points in the scene. The process requires the following steps (see Fig. 2d): First, image features such as SIFT (scale-invariant feature transform) key-points (Lowe 2004) are detected in each image. The features are matched across images, and matching feature points are then used to find the epipolar geometry between pairs of images. Next, in an incremental process, starting from an image pair and adding one image at a time, the feature matches are used to compute a consistent set of camera geometries and 3D scene points (Bundle Adjustment, Snavely et al. 2006). The outcome of this process is a sparse point cloud. In a final step, a dense point cloud is produced from the registered overlapping images by multi-view stereo reconstruction (Furukawa and Ponce 2010). These steps (and more) are combined in the software called *VisualSFM* (http://ccwu.me/vsfm/, Wu et al. 2011, Wu 2013), which is a free tool for Linux, Mac OS and Windows operating platforms. It combines bundle adjustment and a dense 3D scene reconstruction (Furukawa and Ponce 2010), offers a graphical user interface and many features for optimising computing time and the quality of outcomes. It requires a set of images of a scene as input and computes intrinsic camera parameters (if unknown), camera positions and orientations, as well as a (sparse or dense) 3D point cloud of the recorded scene.

For our model we used a total of 120 UV (1392 x 1040 pixel) and RGB images (4000 x 3000 pixel) of a wasp nest area as input. All images were recorded with hand-held cameras from multiple viewpoints, ensuring good coverage of the scene and sufficient overlap in the images to retrieve depth information. From the resulting dense point cloud, we then computed a triangular mesh using poisson reconstruction (Kazhdan et al. 2006). All further processing - cleaning, texturing and scaling

of the meshes - was done using *MeshLab*, an open-source software for processing and editing 3D point clouds and meshes (http://meshlab.sourceforge.net/). We used the Pix4DMapper software that combines all these steps to build models of ant nest environments and along ant foraging paths.

From RGB to false colour UV models: Since all camera images were taken by hand with two different cameras from arbitrary view points, we could not directly combine the colour channels of the RGB and UV images to generate false colour UV-G-B images. Instead, we first reconstructed a single point cloud and poisson mesh from both colour and UV images (e.g. Fig. 2d). In this way we made sure that all images would be registered to each other in a single coordinate system. We overlaid texture and colour from the RGB images on one copy of this mesh and texture and luminance information from the UV images on a second copy of the mesh. As both meshes are identical in their geometry, we could then edit the vertex colour values of the colour mesh to hold the UV values instead of the red colour channel, which represents wavelengths that are unlikely to be seen by insects and so obtain a UV-green-blue 3D mesh.

Hardware: All computations were carried out on standard computers: DELL Precision and Latitude Notebooks, both with Intel core i7 processors, and 8 GB of RAM, and a stand-alone DELL Precision T3600 work station equipped with an Intel Xeon E5 1620 processor, 8 GB of RAM and customised with a NVIDIA GeForce GTX 660 Graphics card (2GB) for faster SIFT feature matching. Owing especially to the large 12 mega pixel colour images, a minimum of 8GB of RAM are required to work with these meshes. Most processing time, not including dense reconstruction, is consumed by pairwise matching of the images. We tested the time *VisualSfM* needed to process a data set of 98 colour images with a resolution of 4000 x 3000pixels on two different computers: A laptop with 8GB of RAM running 32-bit Linux and a workstation with 16GB of RAM, running 64-bit Linux. Results for workstation/laptop were 12s/77s for matching, 36s/72s for sparse reconstruction and 150min/N/A for dense reconstruction.

Registration of scanner-based and camera-based 3D models with insect paths: In cases where scanner/camera-based and purely camera-based 3D models have to be combined, we manually selected corresponding feature points in both models and estimated the rotation, translation and scale of the camera-based model with respect to the laser/camera-based model using a direct method that is optimal in the least squares sense (e.g. Challis 1995). It was also necessary to adjust the colour balance for the scanner/camera model in order to make the transition between both models smooth. Paths of insects were either recorded using differential GPS (see Narendra et al. 2013a) or a high-speed stereo camera system (CR600 x2, Optronis Kehl, Germany, at 250 fps). The mapping between the different model and path reference systems were determined by manually

selecting feature points in the computer model, finding the corresponding GPS or stereo 3D coordinates and then estimating rotation and translation. To render insect views, a set of six virtual perspective cameras was moved in the computer model along the paths of the insects. The six cameras have the same 3D position and field of view of 95°, but each is oriented differently with their optical axes orthogonal to the six faces of a cube. The camera images were remapped to a single panoramic image with equi-rectangular mapping. Alternatively, more realistic mappings are possible, taking account of the known sampling array of insect compound eyes (Stürzl et al. 2010, see Fig. 13 and www.insectvision.org).

Assessing 3D model views against real panoramic images: To ground-truth our model views, we recorded panoramic scenes at defined locations with a Sony Bloggie camera (MHS-PM5, Sony Corp, Japan) placed on a levelled release platform 15 cm off the ground (see Narendra et al. 2013a, Zeil et al. 2014). Concentric panoramic colour images were converted to monochromatic (8bit grey scale) images[1] and un-warped to rectangular panoramas, measuring 1440 x 177 pixels, corresponding to a field of view of 360^{o} x 45^{o}, with a resolution of 4 pixels/degree, using a custom-written Matlab program. Sun glare and reflection artefacts in the sky were removed by using the colour replacement tool in Corel Photo Paint X5 (Corel Corporation, Ottawa, Canada) to copy adjacent sky patches into the corrupted areas. 8 bit grey scale images were converted to floating point arrays and a 80 x 80 pixel Gaussian filter with σ = FWHM/2.355 pixels (with full width at half maximum (FWHM) set to twice an inter-ommatidial angle of 3) was applied before rotational image difference functions were determined using the Matlab circshift function. For each 1 pixel shift, the pixel differences were calculated between the reference image and the shifted image, resulting in 1440 x 177 values that were squared. For each image shift, we then calculated either the mean squared or the root mean squared pixel difference.

Results

Modelling ant habitats

Background

The motivation for this modelling exercise comes from our work on the navigational knowledge of individual *Myrmecia croslandi* foragers in a small urban park, which is representative of the open

[1] Pixel values of grey scale images were computed from RGB values according to
I = 0.299 R + 0.587 G + 0.114 B.

grassy woodlands in the Canberra region (Fig. 3; Narendra et al. 2013a, Zeil et al. 2014). We have shown that most *M. croslandi* foragers from a nest travel to a nest-specific foraging tree (yellow star in Fig. 3a, see also Jayatilaka et al. 2013a) and when displaced 10 to 15 m away from the nest are able to home directly from all compass directions, even from those directions in which they are very unlikely to have been before (red and white paths in Fig. 3a). Moreover, upon release, the ants are able to detect approximate home-bearings by a fast rotational scanning procedure and within 20 cm of the release point (Narendra et al. 2013a, Zeil et al. 2014). Most interestingly, some ants and in particular those that have been released more than 10m away from the nest initially follow their path integration vector, or walk in a direction half-way between that indicated by the path integrator and the true nest direction (yellow paths in Fig. 3a). There thus appears to be an area around the nest and the normal foraging corridor, in which ants in this particular landscape can use the landmark panorama to know where they are and this information degrades with distance from the nest.

Our aim here is to systematically map the navigational information content in this environment and to investigate the range over which nest-directed snapshots, which ants are likely to memorize during their learning walks close to the nest (Nicholson et al. 1999, Müller and Wehner 2010, Graham et al. 2010, Jayatilaka et al. 2013b) can in principle provide displaced ants with information on a heading direction that would bring them back to the nest (Narendra et al. 2013a, Wystrach et al. 2014a, Dewar et al. 2014). Figure 3 shows an aerial view of the area (Fig. 3a), together with four different views of a 3D model of the park, created with a laser scanner – colour camera combination (Fig. 3b). In this 3D model we first rendered panoramic views along transects corresponding to the release directions of ants at two of the nests we have studied previously and subsequently in a 15 x 15 m area around each nest with a grid-spacing of 0.5 m. Panoramic views contain two types of information that are relevant for navigation (Zeil et al. 2003, Stürzl and Zeil 2007, Zeil 2012, Collett et al. 2013a): (1) Heading direction can be found by monitoring image differences or familiarity between a current view and an oriented reference (for instance a nest-directed view during a learning walk) during rotational scanning (e.g. Baddeley et al. 2011, 2012; Zeil et al. 2014, Wystrach et al. 2014a), which generates a rotational image difference function (rotIDF, Zeil et al. 2003, 2014; Stürzl and Zeil 2007; Narendra et al. 2013a,b). If a rotIDF has a detectable minimum through this process of 'alignment matching' (Collett et al. 2013a), it normally points in the direction in which the reference image was oriented and thus provides information on the heading direction towards the nest if the reference image was aligned with the nest direction (Graham et al. 2010, Baddeley et al. 2011, 2012; Wystrach et al. 2012, 2014a; Dewar et al. 2014). (2) Global image differences also depend on the distance from a reference location (transIDF, Zeil et al. 2003, Stürzl and Zeil 2007) and

thus contain information on the relative position to the goal. Minimizing these image differences through translation by any kind of gradient descent allows an agent to pinpoint the reference location (Zeil et al. 2003, Zeil 2012). In practice, the largest image differences are due to misalignment between current and reference view (Zeil et al. 2003), so that the minimum of the rotIDF has to be found first (as suggested by Cartwright and Collett 1983, 1987), before moving in such a way that the remaining image difference is reduced, which corresponds to the transIDF at that particular location (Zeil et al. 2003, Narendra et al. 2013a).

Comparing real and model views

We begin our analysis with a comparison between panoramic images that were recorded with a camera at two reference locations (top, blue-framed panoramas in Fig. 4a) and panoramic views that were rendered in the 3D model at the equivalent locations (bottom, red-framed panoramas, Fig. 4a). We test the quality of rendered views at two locations in three ways: First, we compare the auto-rotational image difference functions of real and rendered images (Fig. 4a-c); second, we asked whether rotIDFs between real and rendered views have a detectable minimum (Fig. 4d) and third, we compare the range over which views provide navigational guidance (the 'catchment areas of snapshots') by mapping the values of the transIDF using rendered or real reference views (Fig. 4e). See below for details on the latter procedure. We calculated the auto-rotational image difference functions, by sliding each image across itself, at two different settings of Gaussian filters (FWHM 0.25^{o} and 3^{o}, Fig. 4b,c) to mimic ant eye resolution (*M. croslandi* appr. 3^{o}) and to document the effect of low-pass filtering on the information content of panoramic scenes (see also Zeil and Stürzl 2007, Wystrach et al., this volume). We confirm that low-pass filtering makes the rotIDF shallower of both rendered (red curves in Fig. 4c) and real images (blue curves in Fig. 4c). Real and rendered views from the same location are indeed similar enough so that the orientation of one can be determined by the minimum of the rotIDF (Fig. 4d). And finally, the catchment areas of real and rendered views have a very similar shape (Fig. 4e). However, it is important to note that a more detailed investigation of the differences between rendered and real views is needed. For instance, the depth of the rotIDF between real and rendered views is much shallower than the auto image difference functions (compare Fig. 4d with Fig. 4c) and the extent of mismatch between auto rotIDFs depends on the particular scene (compare left and right columns in Fig. 4a-c). These differences may have consequences for modelling the detailed search or scanning strategies to find minima in IDFs using rendered images (due to the possible absence of local minima), but they have no impact on mapping navigational information, as we will do next.

Mapping navigational information content

We first show how image difference functions develop along 20 m transects in eight different compass directions around two nests (Fig. 5). We take a snapshot at the nest as reference image and calculate IDFs for panoramic images rendered every 0.5 m up to 20 m away from the nest (see inset Fig. 5a). The IDF surfaces show (1) that the depth of the rotIDFs (along the 'orientation' x-axis) become shallower with distance from the reference location (the nest), which reflects the gradient of the transIDFs along each transect (along the 'distance' y-axis), (2) that the distance over which there is a detectable 'valley' in the IDF surface is shorter in some directions compared to others and that this differs for different nest locations (compare Fig. 5a and b) and (3) that in some directions, the bearings of valley floors (rotIDF minima) change with distance from the reference location, such that the minima may point close to 90^{o} away from the reference direction (e.g. north-west and north surface in Fig. 5b).

To demonstrate the full predictive potential of this analysis, we comprehensively mapped the home-bearing information available around two nests by comparing nest snapshots with current views in an area of 15 m radius around the two nests. Figure 6 shows this for one snapshot orientation directly above the nest for the case that the insects do not (Fig. 6a) or do have additional information on their compass orientation when comparing snapshots (Fig. 6b). Assuming that ants follow the local slope of the IDF, the observed behaviour of ants released at eight different compass bearings 10 m away from the nest (ant paths from Narendra et al. 2013a) is consistent with the IDF map when compass information is available (Fig. 6b). Without compass, i.e. when the minimum of the rotIDF has to be found by scanning through all possible orientations (Fig. 6a), the resulting IDF map cannot predict the paths of ants from the nest at the left when they were released at the north-west and south-west release stations. At these release locations, the initial paths lie outside the 'catchment area' which demarks the area over which a snapshot provides nest-directed information. In this example we determined catchment borders by eye from colour coded IDF values (indicated by a black contour line in Fig. 6). However, the question is whether ants can identify their home direction simply by looking around, as they do, without probing the transIDF gradient with significant translations (Narendra et al. 2013a, Zeil et al. 2014). Unless ants are able to employ some kind of predictive approach (*sensu* Möller 2012), this requires that they are able to select the snapshot orientation that is appropriate to the bearing at which they are released. We address this below by asking how far away from the nest nest-directed snapshots need to be acquired to explain that ants are able to determine home direction from all compass directions, 10-15 m away from the nest. Considering the absence of detailed analyses of the learning walks of ants it is important to note,

however, that ants may store multiple snapshots at or close to the nest in different orientations (as suggested by Mangan and Webb 2009; Möller 2012), and not just nest-directed views (as suggested by Müller and Wehner 2010 and Graham et al. 2010) during learning, which would reduce the amount of scanning needed when computing IDFs. For example, instead of storing just four nest-directed views pointing north, west, south and east, at four positions south, east, north and west of the nest, ants could memorise multiple views with different orientations at each of the four locations, which could in principle be tagged with the direction of the nest.

The range over which views provide guidance

This analysis now allows us to ask firstly, how the navigational information content in this environment, as measured by the range over which panoramic image differences provide effective guidance, depends on the three dimensional layout of landmarks, such as trees and secondly, how the range over which such views can provide guidance to ants released in locations they have not visited before depends on where ants may have acquired views during their learning walks (for an *in silico* analysis of these questions see Dewar et al. 2014). To tackle the first question, we determined the transIDF around one of the nests before (Fig. S1a) and after manually removing in the 3D model two of the closest trees using a suitable software tool for point cloud manipulation such as Meshlab (Fig. S1b). The result confirms our previous analysis (Stürzl and Zeil 2007) that IDFs are narrower and steeper in the presence of close landmarks and become shallower and wider in more open habitat.

Regarding the second question, we had previously predicted that an explanation of the ants' multi-directional homing abilities in this particular landscape would require them to have learnt nest-directed snapshots at distances between 1 and 5 m from the nest (Narendra et al. 2013a). The bearing maps based on nest-directed snapshots shown in Fig. 7 and S2 now allow us to conclude that learning walks extending to between 1.5 and 2 m from the nest would be sufficient to explain the ants' ability to identify the nest direction at locations up to 15 m away from the nest. In the maps shown in Fig. 7, green arrows point in the direction associated with the best matching nest oriented snapshot (the minimum of the rotational IDF between the panoramic image at (x,y) and the best matching snapshot i). For Fig. S2 we assume that compass information is available and that just four IDF values at each position have to be calculated: at each position the current view is aligned with the four snapshots and the image difference is computed. Green arrows show the orientation vector associated with the best matching snapshot and blue arrows point to the direction of the weighted mean vector at each position (following Dewar et al. 2014). However, under the assumption that ants do not know their absolute compass bearing when comparing images, there

remain large areas where the minima of IDFs point away from the true home direction (marked red in Fig. 7). The use of an external compass reference clearly improves this situation (Fig. S2).

We have so far investigated the navigational information provided by the wider landmark panorama in one particular habitat, but had to ignore the complex, fine-scale topography of the ground, which foraging ants are routinely confronted with (e.g. Fig. 8a). In particular for ants displaced to locations they have never been before, navigating through this complex 'undergrowth' must introduce significant visual noise into the process of visual homing, the severity of which needs to be understood. As a first step, we have started to tackle the problem of reconstructing these ground features by using image series recorded with hand-held cameras as input to the camera-based modelling tools described in the Method section. Results are promising and are shown in Fig. 8b-e for the example of the immediate environment of the ant nest photographed in Fig. 8a and for a 3 m stretch of ground leading away from the nest to the foraging tree in Fig. 8f. To tackle the noise problem, these detailed ground-models will need to be embedded into models of the wider landmark panorama in order to reconstruct views from the perspective of ants. We present next an example of such model integration in the more tractable situation of the views encountered by ground-nesting wasps, for which the visual details around the nest are not noise, but provide important guidance.

Modelling views from the cockpit of homing wasps

We employed a hybrid approach with a combination of laser-scans and camera-based methods to reconstruct the views experienced by ground-nesting wasps during their learning and subsequent homing flights (for an earlier attempt see Zeil et al. 2007). A photograph of the nest area is shown in Fig. 9a. The local panorama as raw reflectivity values of a laser scan is shown in Fig. 2b and the point cloud rendered with colour camera information in Fig. 2c. Due to the near-field limitations of the laser scanner, which has a minimum operating range of about 0.5m, the laser-based model of the ground texture becomes increasingly noisy when views are rendered close to the ground (Fig. 9b). We therefore combined laser- and camera-based models (for details see Methods) to arrive at a high-resolution reconstruction of both panorama and ground (Fig. 9c, Fig. 10a). Using calibration markers on the ground we registered the 3D coordinates of wasp flight paths with this model, which we recorded with a high-speed stereo camera system (Fig. 10a) and are thus able to render the views experienced by the insects throughout learning and homing flights. We document the quality and analytical power of this procedure with the example of a learning flight by a wasp (*Cerceris australis*) that occupied the nest marked by a red circle in Fig. 9a. During learning flights, ground-nesting wasps typically fly along ever increasing arcs around the nest while gaining height above

ground at about the same rate as their distance from the nest increases. This results in a cone-shaped flight path, centred on the nest (Fig. 10a). Wasps periodically change pivoting direction (see black line in Fig. 10b) and as they fly along an arc, counterturn in such a way (red line, Fig. 10b) that the nest entrance is seen at lateral retinal positions in the left or right visual field (green line, Fig. 10b; see also Zeil 1993, Zeil et al. 1996, 2007, 2009). The exquisite timing of these flights and the resulting sequence with which a wasp encounters and re-encounters different views is most clearly documented by the matrix of view differences experienced during a learning flight (Fig. 10c).

Fig. 10d shows a sequence of views encountered by the wasp at moments shortly after she reverses pivoting direction (marked by purple spheres in Fig. 10a, blue and red crosses along x-axis in Fig. 10b) and faces the nest entrance. Note that the distant landmark panorama alternatively looks very similar in this sequence of views (compare blue and red framed image pairs in Fig. 10d).This is because the wasp tends to face in the same direction when reversing pivoting direction on the right (blue frames) or the left side of the nest (red frames). Foreground features, however, change from one turning point to the next, because the distance of the wasp from the nest and her height above ground continuously increase.

As views can be rendered at any location within the range of such models, they can be used to test different flight control and homing algorithms in the same complex natural environment. To our knowledge there are few test environments of such complexity that can serve as a benchmark. As one example, we recently addressed the question of how wasps may be able to keep track of their nest entrance during their learning flights (green line Fig. 10b). The problem being that the visual features characterizing the nest entrance change throughout a learning flight (Fig. 10d) due to the wasp's continuous change in perspective and continuously increasing distance to the nest. Using rendered, insect sampling array views (Stürzl et al. 2010) it can be shown, however, that a simple template tracking algorithm with continuous template updating can reliable keep track of the nest entrance location (Fig. 11; Samet et al. 2014). Initially, a template of the nest entrance and its surroundings is extracted from the first insect view of the learning flight and then continuously tracked by searching for the best match (minimum of mean squared pixel difference) between template and the current view within a rectangular area. The search area is centred at the best matching position in the previous template (size 28 x 28 pixels). The template is updated every 5th frame using the best found match.

Discussion

We argue here that there is a need to develop the tools for systematically quantifying navigational information in natural habitats and we have presented the first results of what can be achieved using various methods of rendering panoramic views in 3D models of such environments. The main advantage of our approach compared to previous ones (e.g. Basten and Mallot 2010; Mangan 2011) is the close to veridical view reconstruction it provides. We have shown how the views rendered in 3D models of natural environments can be used to test homing algorithms, such as guidance by view similarities (Fig. 7 and S2, following Graham et al. 2010, Baddeley et al. 2011, 2012), or models of flight control, such as possible mechanisms for tracking of the nest entrance during learning flights (Fig. 11, Samet et al. 2014). These results now form the basis for a suite of specific predictions and experimental tests. For instance, ants can be displaced to areas, which according to our analysis offer no navigational guidance, or we can predict and test the range over which ant views encountered during recorded learning walks can provide navigational information. However, given the novelty of the methods we described here, we will focus our discussion on the strengths and weaknesses of these methods and the opportunities they offer.

Laser scanner based reconstruction

Laser scanners return large data sets (for the Z+F IMAGER 5006i about 10,000 pixel/360° in the default resolution 'high' resulting in about $50x10^6$ depth measurements, recorded in about 10 min) from a single view point. Compared to camera-based 3D model acquisition (see below) depth accuracy is high even for quite distant objects (the maximum range of the Z+F IMAGER 5006i is about 80m), full panoramic 3D acquisition is guaranteed and missing information below the scanner, but also occlusions, can be filled in by scanning the scene from multiple locations. There is a trade-off, however, between range and general noise levels. Since reflections of the laser beam at distant objects usually have low intensity, a low threshold has to be used to accommodate distant objects. On the other hand, a low threshold increases reconstruction noise because sky regions or objects at distances beyond the maximum range will, in particular for phase-based laser scanners, result in erroneously small and low intensity distance measurements. In natural environments (in contrast to indoor scenes), a moderately high threshold is therefore needed to remove such 'infinity' noise. Noisy data points, together with those generated by partial reflection at object boundaries need to be manually detected and removed using Z+F LaserControl software. It is important to realize, however, that distant, visible features, such as mountains play an important role in shaping the range over which rotIDFs can provide bearing or visual compass information (see for instance Towne and Moscrip 2008; Pahl et al. 2011): the more dominant and persistent across different

locations their contribution to the panorama, the higher their navigational information content. When reconstructing this information content in natural scenes, in particular when using laser scans, it is thus crucial to find ways of including these distant features that are lost in the laser scanner voxel cloud, but are present in the camera-based representation of the scene.

The drawbacks of scanner-based reconstruction are the need for additional acquisition and mapping of colour information and the limited number of viewpoints that can be obtained within a reasonable amount of time. Sufficiently many viewpoints are necessary for the reconstruction of cluttered natural scenes containing vegetation with complex structure. The acquisition of colour information is also not trivial. Ideally, images should be recorded from the same viewpoint as the 3D scan, because colour-mapping is then straight-forward once a panoramic image has been created. Our scanner was equipped with a motorized rotating camera (Z+F M-Cam) that starts capturing images immediately after scanning and takes about 3 min to capture the full panorama. However, off-line colour mapping is complex since images and laser scans are recorded from different viewpoints. To remedy this we used in some cases a camera mounted on a nodal point adaptor (ensuring that the camera is rotated around its centre of projection – the “nodal point” of the lens – so that images can be taken in different directions but from the same 3D position), which has to be placed on the tripod after removing the scanner as soon as a scan is completed, an additional step that can take a significant amount of time in the field. In the present context the most serious drawback of laser scanners is the limited number of viewpoints that can be acquired within a reasonable amount of time and the fact that they have a minimum operating range typically between 0.5 and 1m. This does not allow the detailed topography of the ground to be resolved and reconstructed which is especially relevant to walking insects, but also to flying insects when pinpointing goals. It is for this reason that we experimented in addition with camera-based reconstruction methods, which we will discuss next.

Structure from Motion - some recommendations on camera-based reconstruction

In practice, we found that camera-based scene reconstruction of natural environments is unexpectedly complicated. The main reasons are the lack of instantaneous feedback on the quality of reconstruction in current bundle adjustment software and the large number of images that need to be acquired. It is important during acquisition to cover the whole sphere, to ensure sufficient overlap between images for subsequent feature matching, and to record images from many different viewpoints, so that the distance of both close objects (that need small baselines between camera images) and of distant objects (that require large baselines) can be reconstructed. In the future, some of these issues will be less severe thanks to support by bundle adjustment software of

ultra-wide field of view cameras (with FOV around 180°) and the development of tools providing real-time feedback on laptops or even smart phones (e.g. Engel et al. 2014).

The advantage of camera-based reconstruction is clearly that it can make use of any modern digital camera, because small lens distortions as they exist in off-the-shelf consumer cameras can be estimated during bundle adjustment. In the presence of large lens distortions, as they exist in wide angle lenses, the quality of reconstruction can be improved significantly by calibrating the camera and un-distorting the images prior to Structure from Motion (SfM) processing. Using the software *VisualSFM*, it is advisable to calibrate with only one radial distortion parameter and to feed the calibration parameters into the program. If cameras are to be calibrated, it is important to use fixed focus mode. Omnidirectional lenses (panoramic images) can be used for SfM, but are not supported by *VisualSFM*'s bundle adjustment, which supports only a planar pinhole camera model. For omnidirectional imagery, or when it is important to have the ability to define geometric camera constraints (like stereo rigs), a more general purpose bundle adjustment software such as the open source *Ceres-solver* (Agarwal and Mierle 2012) can be used. Video cameras are an option, but care needs to be taken to minimize motion blur by setting fast shutter speeds and good frames need to be selected for SfM, because the quality of the reconstruction is less dependent on the resolution, but rather on the quality of images, that is their 'sharpness', the lens distortions, the textured image content, and the positions relative to each other from which images are recorded. Changing lighting conditions will not significantly affect the reconstruction process itself, but should be taken into account, as shadows may affect the appearance of the final 3D model.

In contrast to the employment of platforms such as robotic arms, cars or UAVs to record images, it is difficult to plan optimal viewpoints using a hand-held camera. In general there should be more than 50% overlap between successive images and rotations should be kept to a minimum to ensure good feature matching. Image features consist of 'points of interest', typically characterised by strong intensity changes, and a photometric descriptor of the region around this point. These regions will change their appearance when seen from different directions and distances. Ideally, feature descriptors should be invariant to changes in rotation and scale, and also in illumination. SIFT features have very good invariance to changes in illumination. They are also scale-invariant, so that changing the distance of a camera to the object does not pose a problem and they are invariant to 2D rotations. In practice, when images are recorded with a hand-held camera, feature matching will only work reliably for viewpoint changes of up to approximately 30^{o}.

Outlook & Opportunities: Towards quantifying navigational information in natural habitats.

We are still far away from being able to reconstruct the full information content of images as they could in principle be perceived by insect eyes. For a start, we lack the tools to render our 3D models with potentially relevant spectral and polarization information, considering that most insects, including ants (Ogawa et al. submitted) are sensitive to the UV, blue and green part of the spectrum and to the direction of polarized light. Both spectral and polarization properties of natural light carry navigation relevant information (e.g. Wehner and Labhart 2006, Möller 2002, Stone et al. 2014). Our models also lack the illumination dynamics of natural scenes, the change in direction of illumination due to the movement of the sun, the predominance of shadows depending on the movement of clouds and the effects of environmental motion, such as wind-driven movement of vegetation. There is also a need to improve the tools for reconstructing viewpoints close to the ground, where many insects, such as ants preform their navigational feats. Finally, apart from a few examples (e.g. Dahmen 1991, Petrovitz et al. 2000, Smolka and Hemmi 2009, Stürzl et al. 2010) we do not have accurate information on the sampling arrays of different insect eyes, have very limited information on early visual processing under natural conditions and on the representation of navigation-relevant information at higher levels of processing in the insect brain (see, however, Homberg et al. 2011, Heinze et al. 2013, Seelig and Jayaraman 2013).

This said, we believe that quantifying navigational information in the natural world will be crucial for testing the validity of models and for solving some of the contentious issues currently being discussed in the animal navigation literature, such as evidence for or against a 'cognitive map' (e.g. Cheung et al. 2014, Cheeseman et al. 2014), or visual *versus* olfactory navigation (e.g. Phillips and Jorge 2014, Wallraff 2014). As we have shown, mapping the navigational information potentially available to animals is beginning to become possible for visual, and, we should add, magnetic navigation (e.g. Boström et al. 2012). 3D models of natural navigation environments can now not only be used to test models of animal navigation under real life conditions, but also can serve as complex natural benchmark environments for critically comparing and testing control and navigation algorithms for outdoor robotic platforms (e.g. Vardy and Möller 2005).

Acknowledgements.

We acknowledge financial support from the Go8 Australia Germany Joint Research Cooperation Scheme, the German Aerospace Centre (DLR), the Australian Research Council (ARC) Centre of Excellence Scheme (CE0561903) and an ARC Discovery Early Career Award (DE120100019). Aerial photographs were supplied by the ACT Planning Authority (ACTPLA).

References

Agarwal S, Mierle K (2012) Ceres Solver. http://ceres-solver.org

Baddeley B, Graham P, Philippides A, Husbands P (2011) Holistic visual encoding of ant-like routes: Navigation without waypoints. Adaptive Behav 19: 3-15

Baddeley B, Graham P, Husbands P, Philippides A (2012) A model of ant route navigation driven by scene familiarity. PLoS Comput Biol 8(1): e1002336. doi:10.1371/journal.pcbi.1002336

Basten K, Mallot HA (2010) Simulated visual homing in desert ant natural environments: efficiency of skyline cues. Biol Cybern 102: 413-425

Boström JE, Åkesson S, Alerstam T (2012) Where on earth can animals use a geomagnetic bi-coordinate map for navigation? Ecography 35: 1039–1047

Briscoe AD, Chittka L (2001) The evolution of color vision in insects. Annu Rev Entomol 46: 471–510

Buehlmann C, Cheng K, Wehner R (2011) Vector-based and landmark guided navigation in desert ants inhabiting landmark-free and landmark-rich environments. J Exp Biol. 214:2845–2853

Buehlmann C, Hansson BS, Knaden M (2012) Path integration controls nest-plume following in desert ants. Curr Biol 22: 645-649

Cartwright BA, Collett TS (1983) Landmark learning in bees: experiments and models. J Comp Physiol 151: 521-543

Cartwright BA, Collett TS (1987) Landmark maps for honeybees. Biol Cybern 57: 85–93

Challis JH (1995) A procedure for determining rigid body transformation parameters. J Biomech 28: 733-737

Cheeseman JF, Millar CD, Greggers U, Lehmann K, Pawley MDM, Gallistel CR, Warman GR, Menzel R (2014) Reply to Cheung et al.: The cognitive map hypothesis remains the best interpretation of the data in honeybee navigation. Proc Nat Acad Sci USA 111: E4398 (doi:10.1073/pnas.1415738111)

Cheng K, Middleton EJT, Wehner R (2012) Vector-based and landmark-guided navigation in desert ants of the same species inhabiting landmark-free and landmark-rich environments. J Exp Biol 215: 3169–3174

Cheung A, Stürzl W, Zeil J, Cheng K (2008) The information content of panoramic images: II. View-based navigation in non-rectangular experimental arenas. J Exp Psychol: Anim Behav Proc 34: 15-30

Cheung A, Collett M, Collett TS, Dewar A, Dyer A, Graham P, Mangan M, Narendra A, Philippides A, Stürzl W, Webb B, Wystrach A, Zeil J (2014) Still no convincing evidence for cognitive map use by honeybees. Proc Nat Acad Sci USA 111: E4396-E4397 (doi/10.1073/pnas.1413581111).

Collett M, Chittka L, Collett TS (2013a) Spatial memory in insect navigation. Curr Biol 23: R789-R800

Collett TS, Hempel de lbarra N, Riabinina O, Philippides A (2013b) Coordinating compass-based and nest-based flight directions during bumblebee learning and return flights. J Exp Biol 216: 1105-1113

Dahmen HJ (1991) Eye specialisation in waterstriders: an adaptation to life in a flat world. J Comp Physiol A 169: 623-632

Dewar ADM, Philippides A, Graham P (2014) What is the relationship between visual environment and the form of ant learning-walks? An *in silico* investigation of insect navigation. Adaptive Behavior 22: 163–179

Engel J, Schöps, T, Cremers D (2014) LSD-SLAM: Large-scale direct monocular SLAM. In Computer Vision – ECCV 2014. Fleet D, Pajdla T, Schiele B, Tuytelaars T (eds). Lecture Notes in Computer Science 8690: 834-849

Furukawa Y, Ponce J (2010) Accurate, dense, and robust multiview stereopsis. IEEE Trans Pattern Anal Machine Intell 32: 1362–1376

Graham P, Philippides A, Baddeley B (2010) Animal cognition: multi-modal interactions in ant learning. Curr Biol 20: R639-R640

Graham P, Cheng K (2009) Ants use the panoramic skyline as a visual cue during navigation. Curr Biol 19: R935-R937

Graham P, Cheng K (2009) Which portion of the natural panorama is used for view-based navigation in the Australian desert ant? J Comp Physiol A 195: 681-689

Hartley R, Zisserman A (2003) Multiple view geometry in computer vision. Cambridge University Press, Cambridge.

Heinze S, Florman J, Asokaraj S, el Jundi B, Reppert SM (2013) Anatomical basis of sun compass navigation II: The neuronal composition of the central complex of the Monarch Butterfly. J Comp Neurol 521: 267-298

Homberg U, Heinze S, Pfeiffer K, Kinoshita M, el Jundi B (2011) Central neural coding of sky polarization in insects. Phil Trans R Soc B 366: 680-687

Jayatilaka P, Raderschall CA, Narendra A, Zeil J (2013a) Individual foraging patterns of the jack jumper ant, *Myrmecia croslandi*. Myrmecol News 19: 75-83

Jayatilaka P, Raderschall CA, Zeil J, Narendra A (2013b) Learning to forage: the learning walks of Australian jack jumper ants. Front Physiol Conference Abstract: Int Conf Invertebrate Vision. doi: 10.3389/conf.fphys.2013.25.00081

Kazhdan M, Bolitho M, Hoppe H (2006) Poisson surface reconstruction. Proc fourth Eurographics Symposium on Geometry Processing: 61-70

Kohler M, Wehner R (2005) Idiosyncratic route-based memories in desert ants, *Melophorus bagoti*: How do they interact with path-integration vectors? Neurobiol Learn Memory 83: 1–12

Labhart T (1986) The electrophysiology of photoreceptors in different eye regions of the desert ant, *Cataglyphis bicolor*. J Comp Physiol A 158: 1–7

Legge ELG, Wystrach A, Spetch ML, Cheng K (2014) Combining sky and earth: desert ants (*Melophorus bagoti*) show weighted integration of celestial and terrestrial cues. J Exp Biol 217: 4159-4166

Lowe DG (2004) Distinctive image features from scale-invariant keypoints. Int J Comput Vis 60: 91–110

Mair E, Stürzl W, Zeil J (2013) Benchmark 3D models of natural navigation environments @ www.InsectVision.org. Front Physiol Conference Abstract: Int Conf Invertebrate Vision. doi: 10.3389/conf.fphys.2013.25.00084

Mangan M (2011) Visual homing in field crickets and desert ants: a comparative behavioural and modelling study. PhD thesis. School of Informatics, University of Edinburgh

Mangan M, Webb B (2009) Modelling place memory in crickets. Biol Cybern 101:307–323

Mangan M, Webb B (2012) Spontaneous formation of multiple routes in individual desert ants (*Cataglyphis velox*). Behav Ecol 23: 944-954

Möller R (2002) Insects could exploit UV-green contrast for landmark navigation. J Theor Biol 214: 619–631

Möller R (2012) A model of ant navigation based on visual prediction. J Theor Biol 305: 118-130

Müller M, Wehner R (2010) Path integration provides a scaffold for landmark learning in desert ants. Curr Biol 20: 1368–1371

Narendra A (2007a) Homing strategies of the Australian desert ant *Melophorus bagoti* I. Proportional path-integration takes the ant half-way home. J Exp Biol 210: 1798-1803

Narendra A (2007b) Homing strategies of the Australian desert ant *Melophorus bagoti* II. Interaction of the path integrator with visual cue information. J Exp Biol 210: 1804-1812

Narendra A, Gourmaud S, Zeil J (2013a) Mapping the navigational knowledge of individually foraging ants *Myrmecia croslandi*. Proc R Soc Lond B 280: 20130683

Narendra A, Raderschall CA, Robson SKA (2013b) Homing abilities of the Australian intertidal ant, *Polyrhachis sokolova*. J Exp Biol 216: 3674-3681

Nicholson DJ, Judd SPD, Cartwright BA, Collett TS (1999) Learning walks and landmark guidance in wood ants (*Formica rufa*). J Exp Biol 202: 1831-1838

Ogawa Y, Falkowski M, Narendra A, Zeil J, Hemmi JM (2015) Three spectrally distinct photoreceptor types in Australian bull ants. In preparation.

Pahl M, Zhu H, Tautz J, Zhang S (2011) Large scale homing in honeybees. PLoS ONE 6: e19669. doi:10.1371/journal.pone.0019669

Philippides A, Baddeley B, Cheng K, Graham P (2011) How might ants use panoramic views for route navigation? J Exp Biol 214: 445–451

Philippides, A., Hempel de lbarra, N, Riabinina O, Collett TS (2013) Bumblebee calligraphy: the design and control of flight motifs in the learning and return flights of *Bombus terrestris*. J Exp Biol 216: 1093-1104

Petrowitz R, Dahmen H, Egelhaaf M, Krapp HG (2000) Arrangement of optical axes and spatial resolution in the compound eye of the female blowfly *Calliphora*. J Comp Physiol A 186: 737-746

Phillips JB, Jorge PE (2014) Olfactory navigation: failure to attempt replication of critical experiments keeps controversy alive. Reply to Wallraff. Anim Behav 90: e7-e9

Pollefeys M, van Gool L, Vergauwen M, Verbiest F, Cornelis K, Tops J, Koch R (2004) Visual 3D modeling with a hand-held camera. Int J Comp Vis 59: 207-232

Samet N, Zeil J, Mair E, Boeddeker N, Stürzl W (2014) Ground-nesting insects could use visual tracking for monitoring nest position during learning flights. In From Animals to Animats 13, del Pobil AP, Chinellato E, Martínez-Martín E, Hallam J, Cervera E, Morales A (eds). Lecture Notes in Computer Science 8575: 108-120

Sandoval EL, Wajnberg E, Esquivel DMS, Lins de Barros H, Acosta-Avalos D (2012) Magnetic orientation in *Solenopsis sp*. ants. J Insect Behav 25: 612-619

Schultheiss P, Wystrach A, Legge ELG, Cheng K (2013) Information content of visual scenes influences systematic search of desert ants. J Exp Biol 216: 742-749

Seelig JD, Jayaraman V (2013) Feature detection and orientation tuning in the *Drosophila* central complex. Nature 503: 262-266

Sellers WI, Hirasaki E (2014) Markerless 3D motion capture for animal locomotion studies. Biology open 3(7): 656–668

Smolka J, Hemmi JM (2009) Topography of vision and behaviour. J Exp Biol 212: 3522-3532

Snavely N, Seitz SM, Szeliski R (2006) Photo tourism: exploring photo collections in 3D. ACM transactions on graphics (TOG) 25: 835–846

Sommer S, von Beeren C, Wehner R (2008) Multiroute memories in desert ants. Proc Nat Acad Sci USA 105: 317–322

Steck K, Hansson BS, Knaden M (2011) Desert ants benefit from combining visual and olfactory landmarks. J Exp Biol 214: 1307-1312

Stone T, Mangan M, Ardin P, Webb B (2014) Sky segmentation with ultraviolet images can be used for navigation. Proc Robotics: Science and Systems 2014

Stürzl W, Zeil J (2007) Depth, contrast and view-based homing in outdoor scenes. Biol Cybern 96: 519–531

Stürzl W, Cheung A, Cheng K, Zeil J (2008) The information content of panoramic images: I. Rotational errors and the similarity of views in rectangular experimental arenas. J Exp Psychol: Anim Behav Proc 34: 1-14

Stürzl W, Boeddeker N, Dittmar L, Egelhaaf M (2010) Mimicking honeybee eyes with a 280° field of view catadioptric imaging system. Bioinspiration & Biomimetics 5: 036002

Stürzl W, Mair E, Hirschmüller H, Zeil J (2013) Mapping the navigational information content of insect habitats. Front Physiol Conference Abstract: Int Conf Invertebrate Vision. doi: 10.3389/conf.fphys.2013.25.00085

Towne WF, Moscrip H (2008) The connection between landscapes and the solar ephemeris in honeybees. J Exp Biol 211: 3729-3736

Vardy A, Möller R (2005) Biologically plausible visual homing methods based on optical flow techniques. Connection Sci 17: 47-89

von Frisch K, Lindauer M (1954) Himmel und Erde in Konkurrenz bei der Orientierung der Bienen. Naturwiss 41: 245-253.

Wajnberg E, Acosta-Avalos D, Alves OC, Ferreira de Oliveira J, Srygley RB, Esquivel DMS (2010) Magnetoreception in eusocial insects: an update. J R Soc Interface 7: S207–S225 (doi: 10.1098/rsif.2009.0526.focus)

Wallraff HG (2014) Do olfactory stimuli provide positional information for home-oriented avian navigation? Anim Behav 90: e1-e6

Wehner R (1997) The ant's celestial compass system: spectral and polarization channels. In Lehrer M (ed) Orientation and Communication in Arthropods. Birkhäuser Verlag, Basel, pp. 145-185

Wehner R (2008) The desert ant's navigational toolkit: Procedural rather than positional knowledge. Navigation 55: 101-114

Wehner R, Labhart T (2006) Polarization vision. In: Nilsson DE, Warrant EJ (eds) Invertebrate vision. Cambridge University Press, Cambridge, pp. 291–348

Wehner R, Müller M (2006) The significance of direct sunlight and polarized skylight in the ant's celestial system of navigation. Proc Natl Acad Sci USA 103: 12575–12579

Wehner R, Michel B, Antonsen P (1996) Visual navigation in insects: Coupling of egocentric and geocentric information. J Exp Biol 199: 129-140

Wolf H, Wehner R (2000) Pinpointing food sources: Olfactory and anemotactic orientation in desert ants, *Cataglyphis fortis*. J Exp Biol 203: 857-868

Wohlfeil J, Strackenbrock B, Kossyk I (2013) Automated high resolution 3D reconstruction of cultural heritage using multi-scale sensor systems and semi-global matching. Int Arch Photogrammetry, Remote Sensing and Spatial Information Sciences, XL-4(W4): 37-43

Wu C (2013) Towards linear-time incremental structure from motion. In: Proceedings of the International Conference on 3D Vision: 127-134

Wu C, Agarwal S, Curless B, Seitz SM (2011) Multicore bundle adjustment. In: Proceedings of the 24th Conference on Computer Vision and Pattern Recognition: 3057-3064

Wystrach A, Beugnon G (2009) Ants learn geometry and features. Curr Biol 19: 61–66. doi:10.1016/j.cub.2008.11.054

Wystrach A, Graham P (2012) What can we learn from studies of insect navigation? Anim Behav 84: 13-20. (doi:10.1016/j.anbehav.2012.04.017)

Wystrach A, Schwarz S (2013) Ants use a predictive mechanism to compensate for passive displacements by wind. Curr Biol 23: R1083-1085

Wystrach A, Beugnon G, Cheng K (2012) Ants might use different view-matching strategies on and off the route. J Exp Biol 215: 44–55

Wystrach A, Cheng K, Sosa S, Beugnon G (2011) Geometry, features, and panoramic views: ants in rectangular arenas. J Exp Psychol Anim Behav Process 37: 420-435

Wystrach A, Philippides A, Aurejac A, Cheng K, Graham P (2014a) Visual scanning behaviours and their role in the navigation of the Australian desert ant *Melophorus bagoti*. J Comp Pysiol A 200: 615-626

Wystrach A, Schwarz S, Schultheiss P, Baniel A, Cheng K (2014b) Multiple sources of celestial compass information in the Central Australian desert ant *Melophorus bagoti*. J Comp Physiol A 200: 591-601

Zeil J (1993) Orientation flights of solitary wasps (*Cerceris*; Sphecidae; Hymenoptera): I. Description of flight. J Comp Physiol A172: 189-205

Zeil J (2012) Visual homing – An insect perspective. Curr Opin Neurobiol 22: 285–293

Zeil J, Boeddeker N, Stürzl W (2009) Visual homing in insects and robots. In Flying Insects and Robots. Floreano D, Zufferey J-C, Srinivasan MV, Ellington C (eds) Berlin Heidelberg New York: Springer Verlag, pp. 87-100

Zeil J, Hofmann MI, Chahl JS (2003) Catchment areas of panoramic snapshots in outdoor scenes. J Opt Soc Am A 20: 450-469

Zeil J, Kelber A, Voss R (1996) Structure and function of learning flights in bees and wasps. J Exp Biol 199: 245-252

Zeil J, Narendra A, Stürzl W (2014) Looking and homing: How displaced ants decide where to go. Phil Trans R Soc Lond B 369: 20130034

Zeil J, Boeddeker N, Hemmi JM, Stürzl W (2007) Going wild: Toward an ecology of visual information processing. In North G, Greenspan R (eds) Invertebrate Neurobiology. Cold Spring Harbor Press, Cambridge Mass, pp. 381-403

Figure Legends

Figure 1 The route memories of desert ants (*Cataglyphis fortis*) in a landmark-rich habitat of low shrubs (indicated by contour lines in 15 cm height intervals) as shown in Wehner et al. 1996. Two ants (red and green) had been trained to visit a feeder (F) 30 m north-west of their nest. Their homing paths are shown as dotted lines. After returning to the nest, the ants were caught and displaced back to the feeder location. After some searching, both ants practically retraced their steps (solid red and green paths). Modified from Wehner et al. 1996.

Figure 2 Laser scanner and camera-based reconstruction methods. **a** Laser scanner/colour camera combination (Z+F Imager 5006i with Z+F M-cam) mounted and levelled on a tripod. **b** Panorama of laser-scanner reflectivity values as recorded at the wasp nesting site. **c** Same scene after registering colour camera images with the laser scan. **d** Flow chart for purely camera-based reconstruction of detailed ground topography. Photographs and coloured meshes show the nesting area of ground-nesting wasps and bees.

Figure 3 Laser scanner-based reconstruction of landscape-scale navigation habitats. **a** Aerial photograph of an urban grassy woodland showing in addition the paths of ants from one nest that have been caught at their foraging tree (yellow star) and displaced 10 m away from the nest in different compass directions (black circles) as full-vector (red paths) or after they had returned to the nest, as zero-vector ants (white paths). Some full vector ants (yellow paths) were released at locations 15-20 m away from the nest where some of them first followed their path integration vector (modified from Narendra et al. 2013a). **b** Different views of the 3D model of this park created from six laser scans together with transects radiating from two nests (pink and blue dashed lines) along which we reconstructed panoramic views for the analysis shown in Fig. 4 and 5. Bright patches with round 'holes' indicate laser scanner positions.

Figure 4 Comparing real and rendered panoramic views. **a** Panoramic photographs (blue-framed top images) at two sites (left and right) and views rendered in the 3D model at the same sites (red-framed bottom images) shown at a resolution of 0.25^{o}. **b** Same views (photographs top, rendered views bottom) low-pass filtered with a Gaussian with full width at half maximum (FWHM) of twice an assumed interommatidial angle of 3^{o}. **c** Rotational image difference functions (rotIDF) for each of the images shown in **a** and b with colours corresponding to the frames around images

(photographs in blue hues and rendered images in red hues). **d** Rotational image difference functions between rendered and real panoramas at the same two sites. **e** The range over which views provide navigational guidance (the 'catchment areas of snapshots'). The maps show the colour-coded values of the translational image difference function (transIDF) using rendered (left) or real reference views (right). See text for details.

Figure 5 Mapping the navigational information content of habitats I. Surfaces show the rotational image difference functions (rotIDF) at 0.5 m intervals along transects radiating out from two nests (**a** and **b**) between the views at these locations and snapshots at the nest that are oriented opposite to the direction of a given transect (e.g. nest snapshot oriented south for the transect north of the nest). See inset in **a** for explanation of axes: The sums of squared pixel differences (z-axis) are plotted over the orientation of a view relative to the reference snapshot (x-axis) and the distance from the nest (y-axis). The 'length of valleys' leading into the minimum of these surfaces indicates the range over which ants would in principle be able to access information on heading direction towards the nest by detecting the minimum of the rotIDF. Note how the location of the nest and the location of trees in this particular landscape determine the shape of these image difference surfaces.

Figure 6 Mapping the navigational information content of habitats II. The value of the translational image difference function (transIDF) mapped over a 15 m radius around two nests, for the case that **a** all orientations are tested when comparing a view with an oriented nest snapshot, or **b** when assuming that ants know their current orientation and compare the current view to a snapshot with the same orientation, which minimizes the effect of false minima. Black contour lines mark the limits of the catchment area of snapshots, outside of which the snapshot does not provide nest-directed information. Superimposed are the paths of full-vector (red) and zero-vector ants (blue) from Narendra et al. 2013a.

Figure 7 How learning walks affect the range of navigational information I. For this analysis we assumed the acquisition of nest-directed snapshots at distances of 1.5 m **a** and 2 m **b** from four different compass bearings (marked by red crosses). The nest is marked by a black circle. At each position (x,y), *green arrows* point in the direction associated with the best matching nest-oriented snapshot with direction vectors $\mathbf{v}(x,y) = \mathbf{v}_{i^*}$, where $i^* = \arg\min_i \text{minrotIDF}(\text{Img}(x,y), \text{Img}_i)$; $\text{minrotIDF}(\text{Img}(x,y), \text{Img}_i)$ is the minimum of the rotIDF between the panoramic image at (x,y) and

snapshot i. *Blue arrows* are the weighted mean vectors calculated as described in Dewar et al. (2014), with $\mathbf{v}(x,y) = \sum_i w_i \mathbf{v}_i$, where $w_i = \text{minrotIDF}(\text{Img}(x,y), \text{Img}_i) / \text{minrotIDF}(\text{Img}(x,y), \text{Img}_i) \leq 1$ ensures that directions associated with snapshots having higher similarity to the image at position (x,y) receive higher weights. Note that the length of blue vectors decreases with distance to snapshot positions and could be used as a confidence measure. *Red areas* approximately mark regions where both green and blue vectors point away from the nest. Exclusively above horizon panoramas were used for this analysis.

Figure 8 **Fine-scale modelling of ground structures using** image series recorded with a hand-held camera. **a** Photograph of the ground around a nest (*red circle*) of the jack jumper ant *Myrmecia croslandi*. **b-e** Views generated in the model of this ground patch with view-points closer and closer to the ground. **f** Top-down view of a model covering a 3 m stretch along the paths of ants in the direction of their nest-specific foraging tree. Models and model views generated and processed with Pix4DMapper by Pix4D (Lausanne, Switzerland).

Figure 9 Combining detailed camera-based and Laser-scanner reconstruction for 3D models including detailed topography on the ground. **a** Photograph of a ground-nesting wasp nesting area with one nest location marked by a red circle. **b** Close-up ground features of the 3D model derived from laser scanner/colour camera combination illustrating the limits of close-range laser scanner data. **c** Combining laser scanner- and purely camera-based reconstruction of close-up ground features, demonstrating the high quality of 3D modelling that can be achieved. Insets show different perspectives of the same scene.

Figure 10 Reconstructing the view from the cockpit of a learning wasp. **a** Two views of a learning flight seen against the 3D model of the ground. The position of the wasp is marked by blue dots every 8 ms. Purple spheres indicate the positions of the panoramic snapshots shown in **d**. **b** The time course for the same learning flight of gaze direction (red), bearing (black), retinal azimuth position of the nest entrance (green), retinal elevation of the nest entrance (light blue) and height above ground (dark blue). The moments at which snapshots in **d** were rendered are marked by blue and red crosses along the dashed zero-degree line. **c** The auto-image difference matrix for this learning flight. The false colour code shows the difference between each view with all other views in the sequence in which they are encountered during the flight. Dashed blue and red lines mark the

times at which the views shown in **d** were rendered. **d** A series of rendered panoramas seen by the learning wasp at the locations and times indicated in **a**, **b** and **c**, in moments when she reversed pivoting direction on the right (blue frames) or the left side of the nest (red frames). Images were created from six (3x2) virtual camera views, each covering 95°x95°, by means of bilinear remapping.

Figure 11 Template-based tracking of the nest entrance during a learning flight. *Left column*: sample panoramic images as seen by the wasp early **a,b** and late **c,d** during the learning flight sequence shown in Fig. 10. *Centre column*: the same views as they would be represented by an insect's sampling array. Each facet is displayed as 2x2 pixels (see Stürzl et al. 2010). *Right column*: insect sampling array views as grey level images with template tracking results shown as the estimated nest position (green dot), which is defined as the centre of the area that provides the best match with the current template (green square, size 22x22 pixels) within the search region (blue square, size 28x28 pixels). True nest position is indicated by the red dot.

Figure
Click here to download Figure: Stuerzl et al_Figures_11-03-2015.pdf

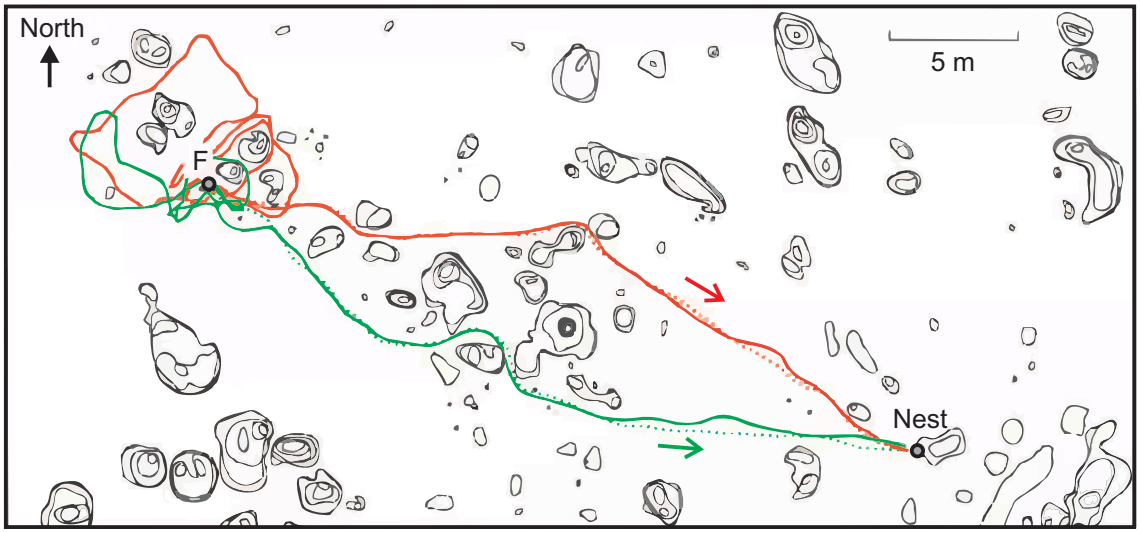


Fig. 1
Stuerzl et al

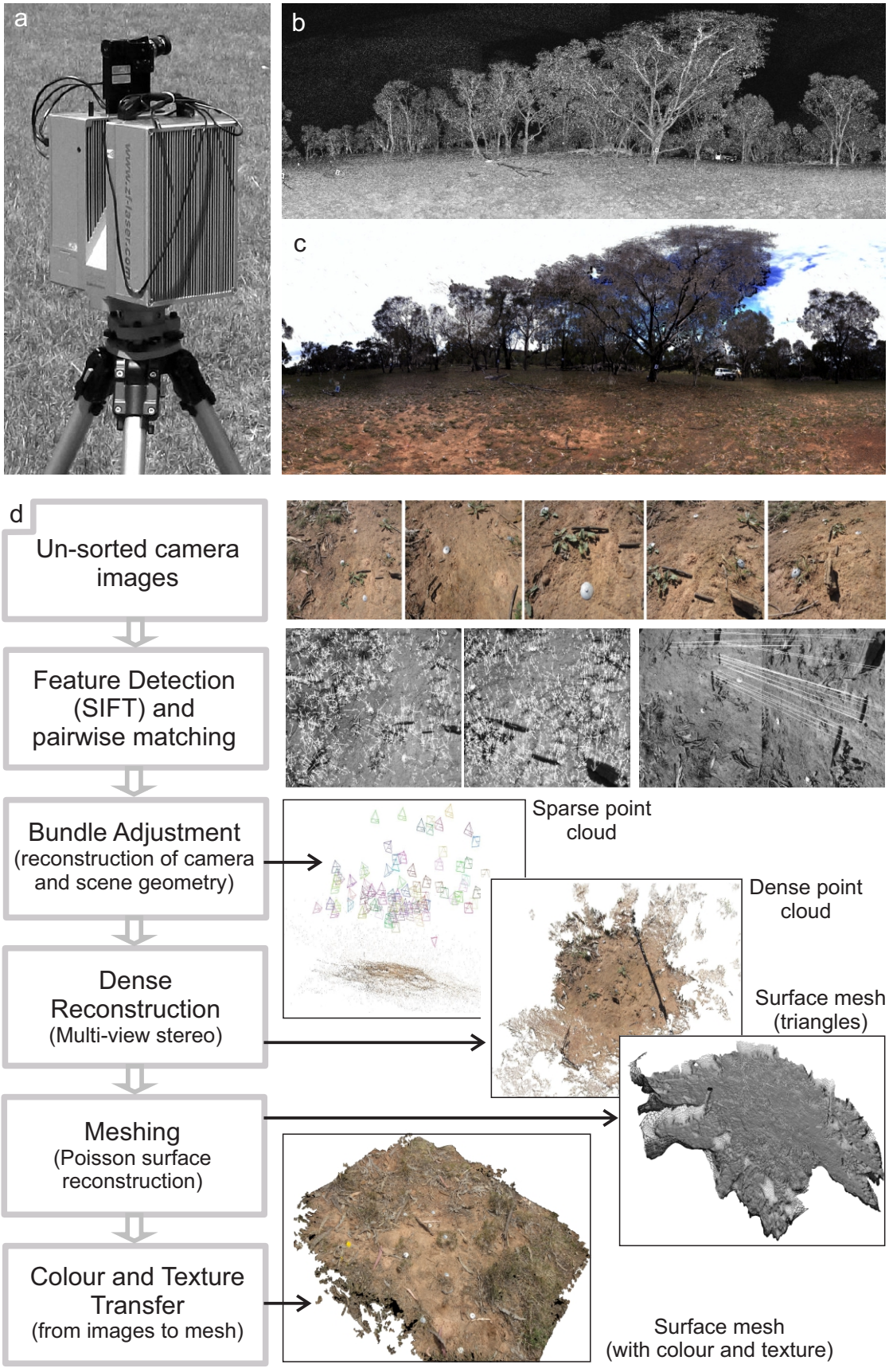


Fig. 2
Stuerzl et al

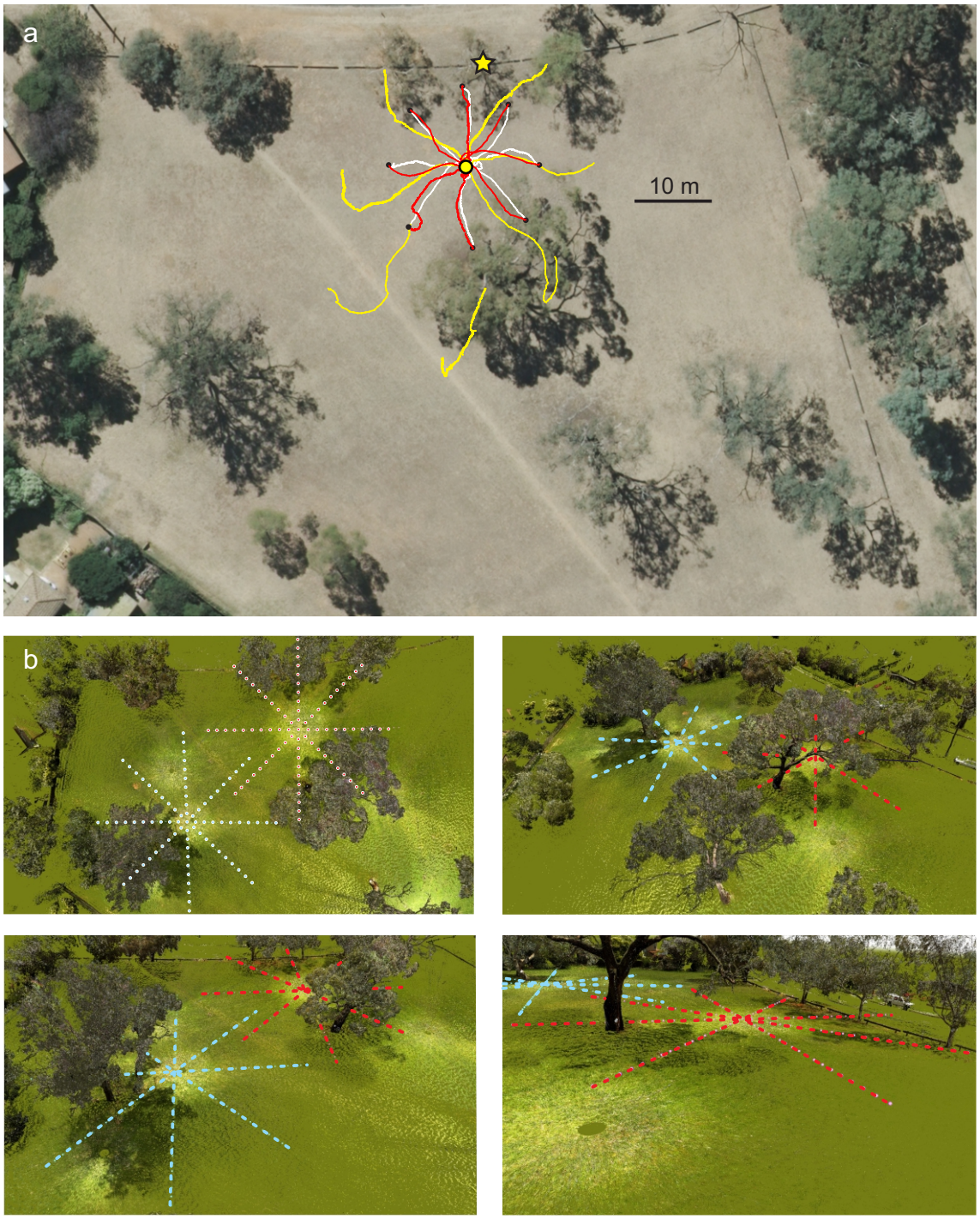


Fig. 3
Stuerzl et al

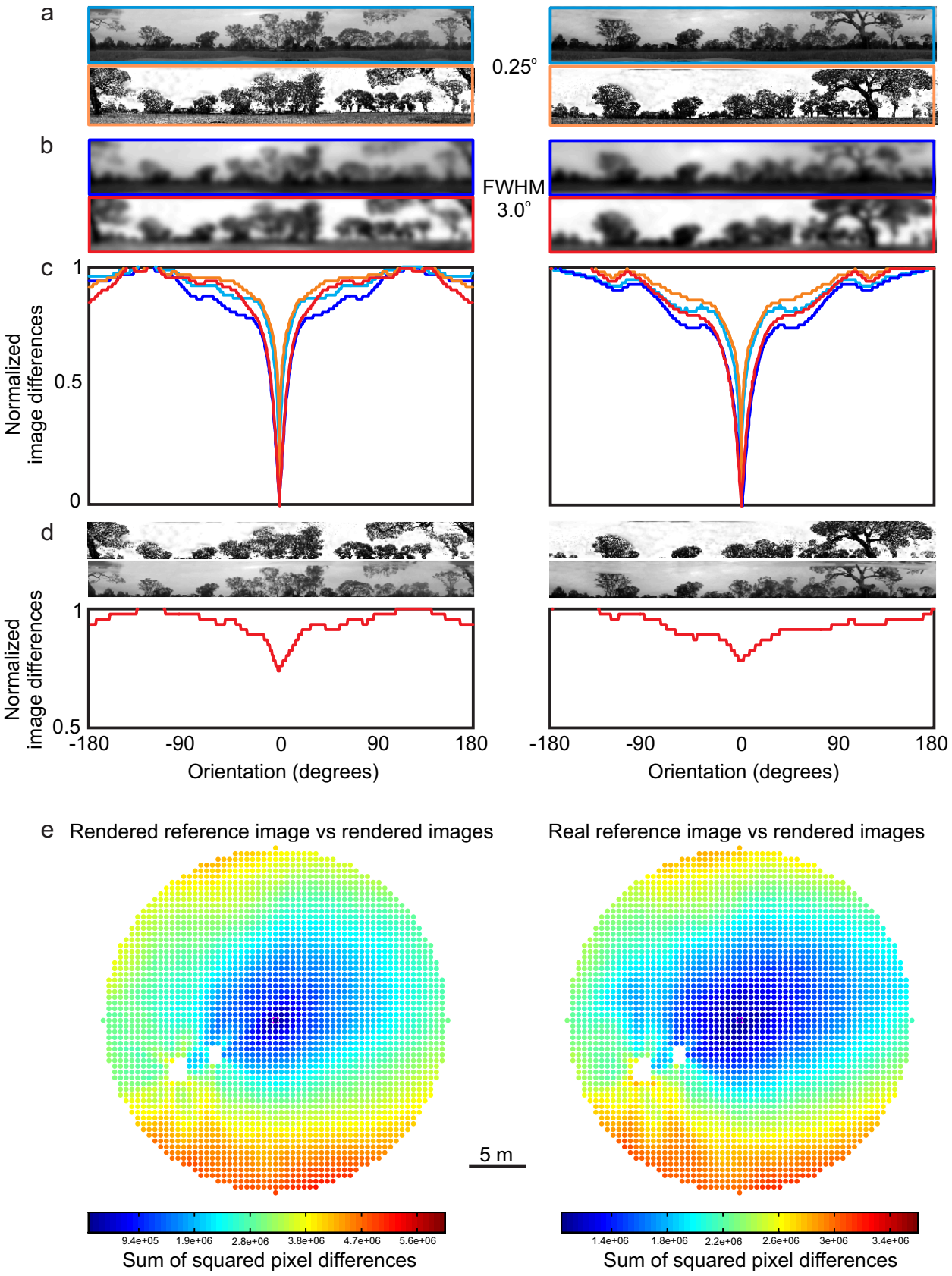


Fig. 4
Stuerzl et al

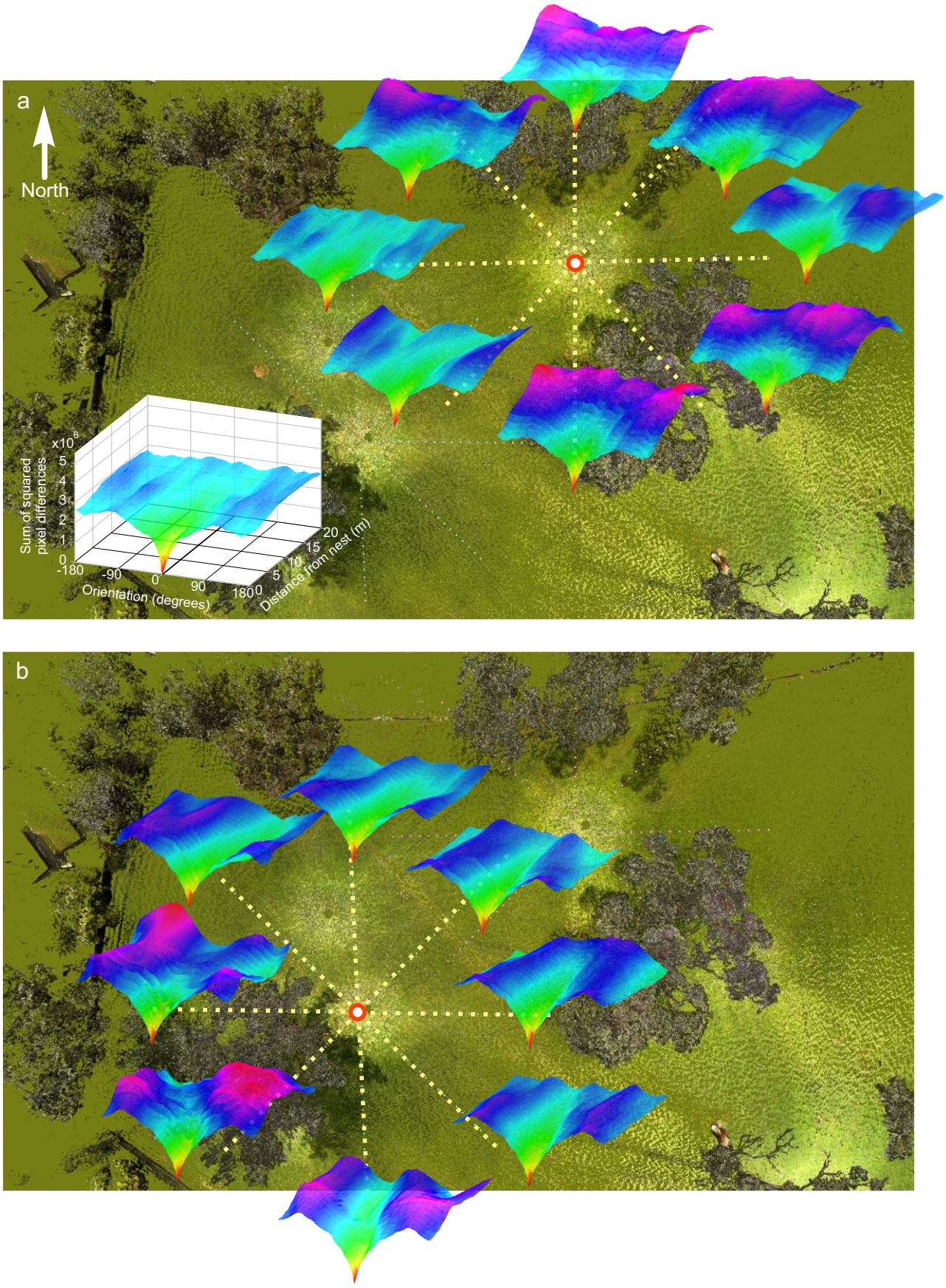
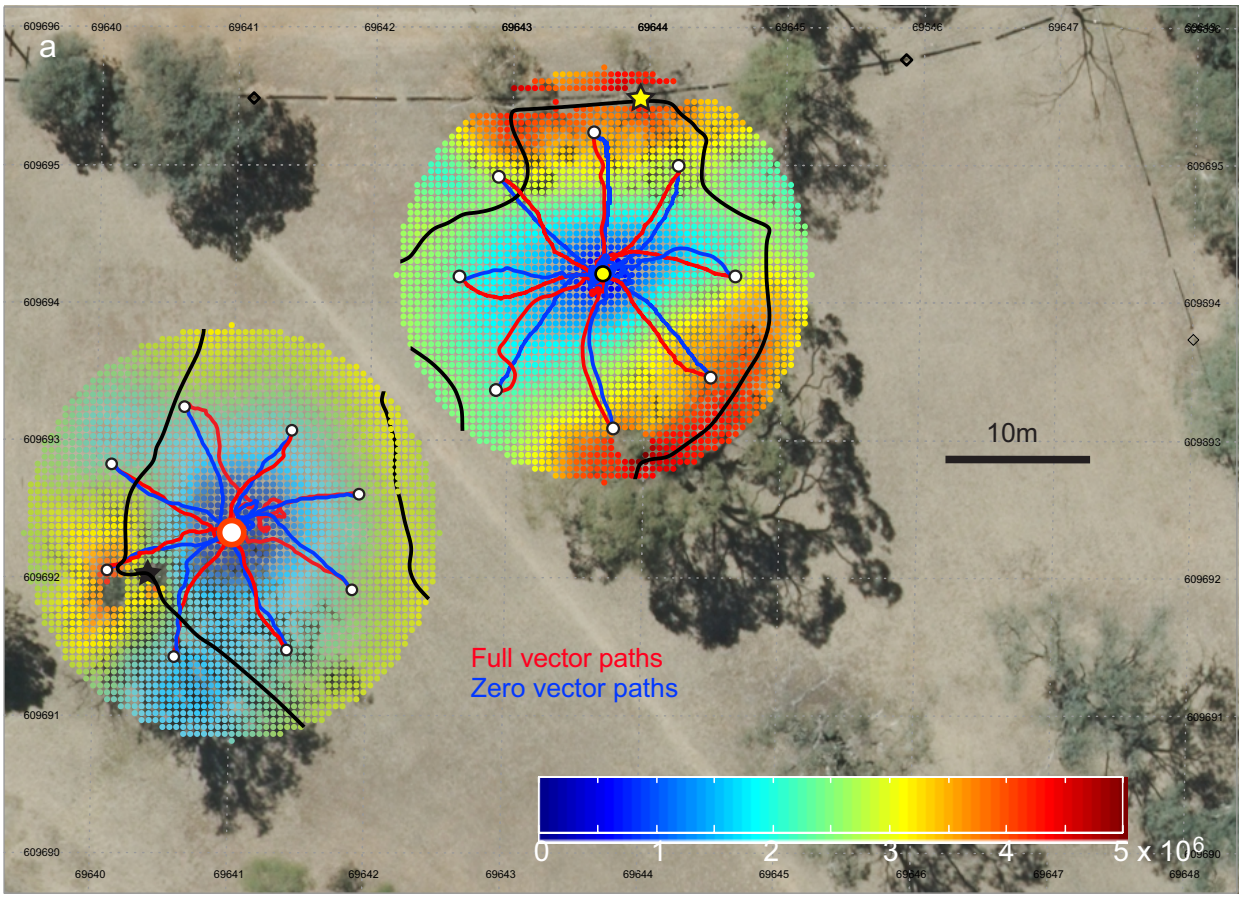


Fig. 5
Stuerzl et al



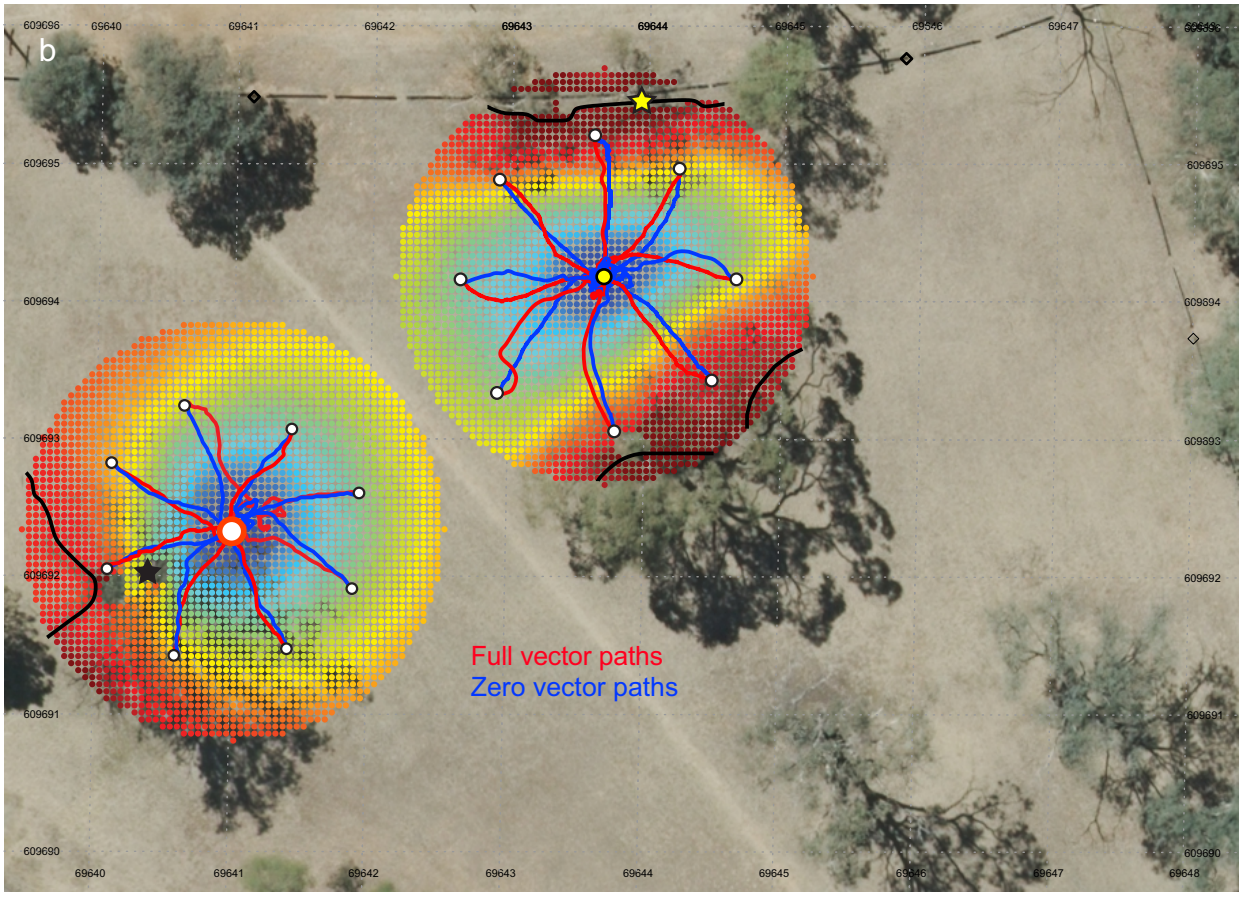
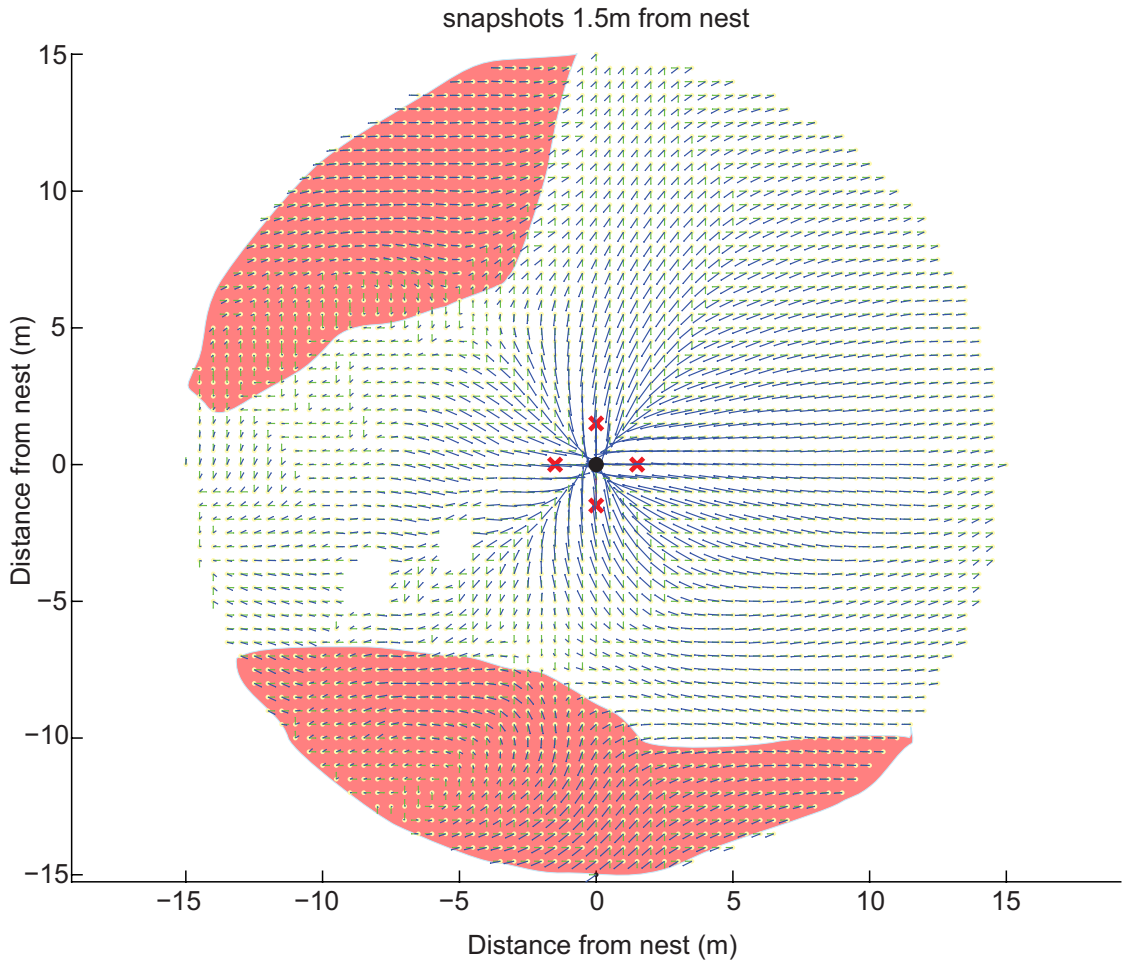


Fig. 6
Stuerzl et al



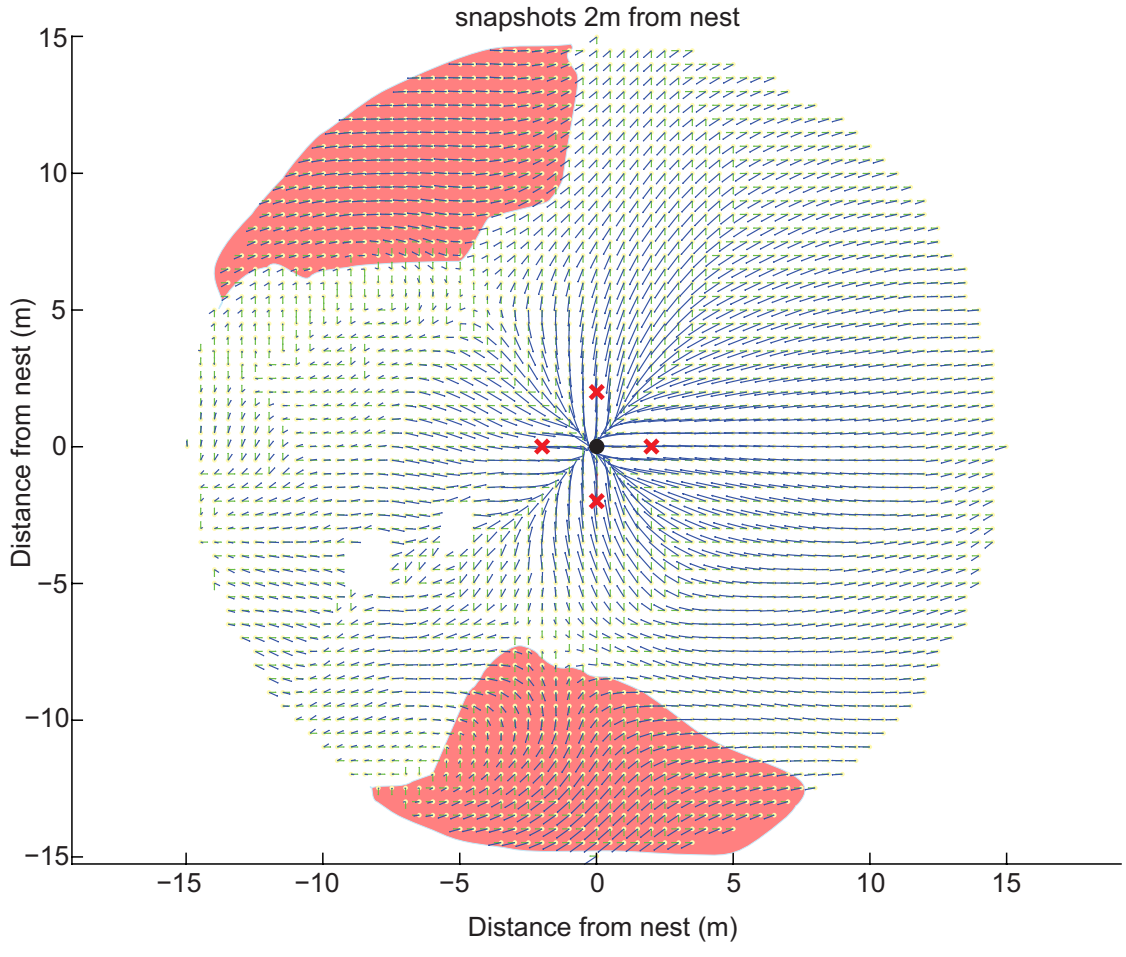


Fig. 7
Stuerzl et al

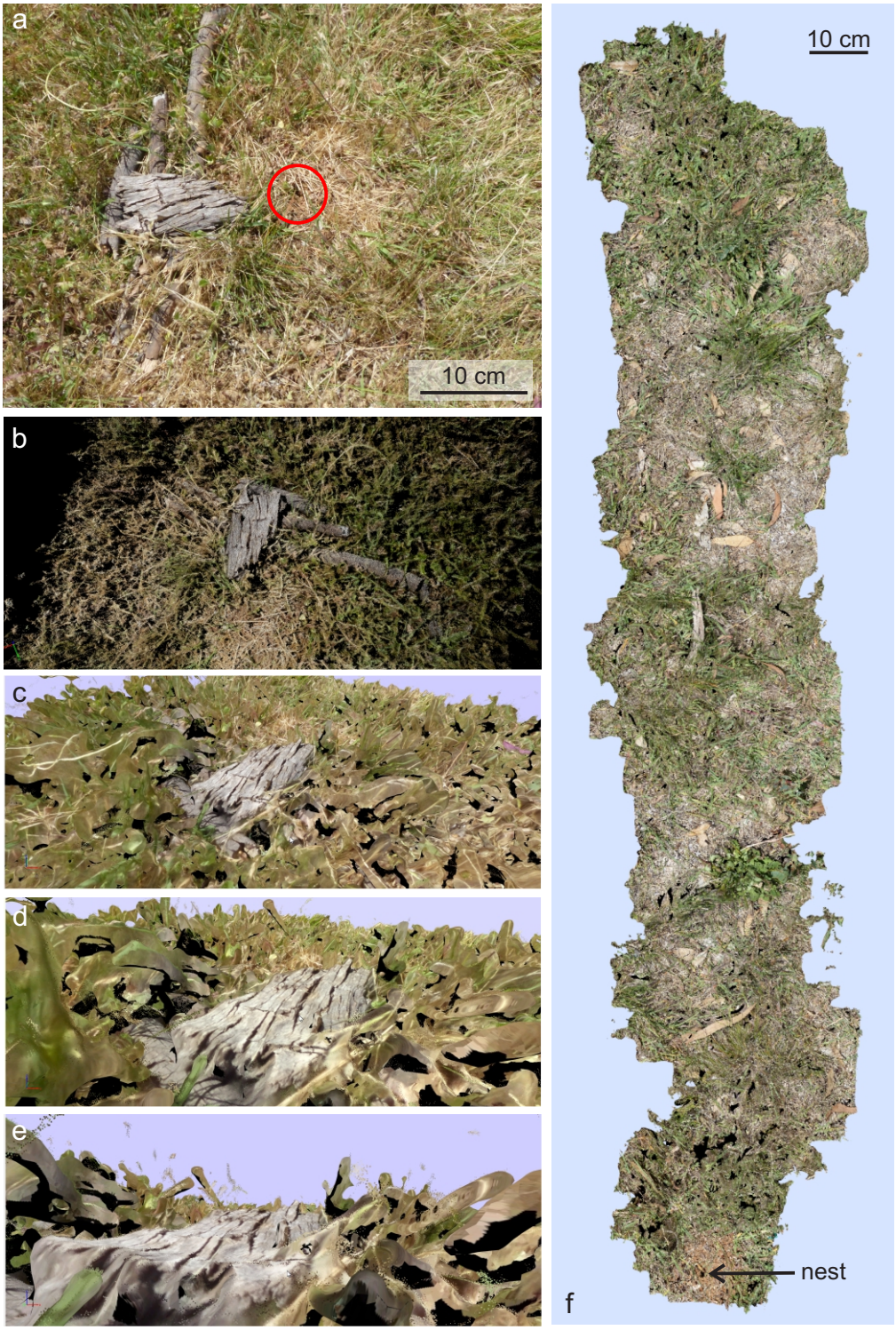


Fig. 8
Stuerzl et al

Fig. 9
Stuerzl et al

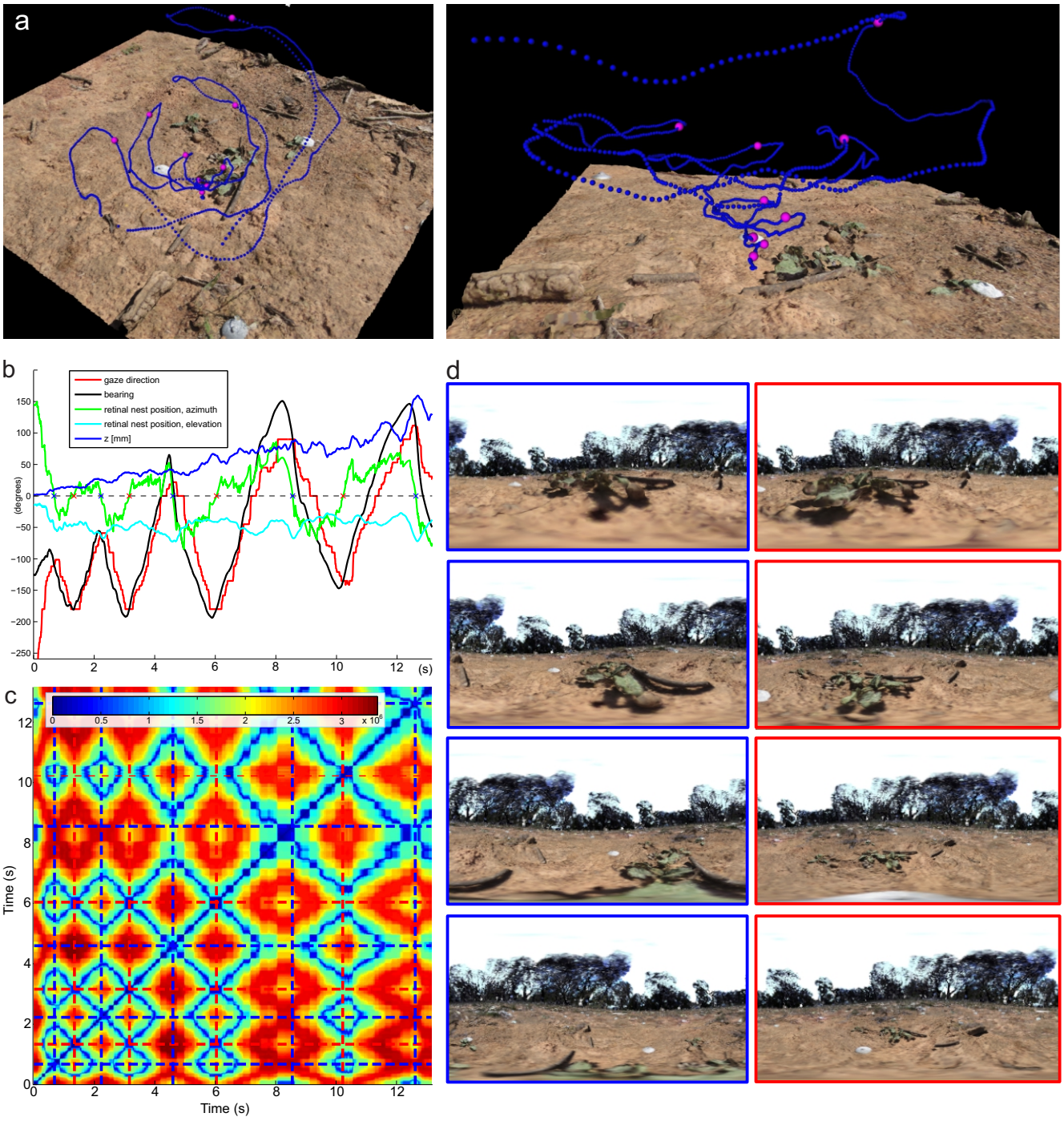


Fig.10
Stuerzl et al

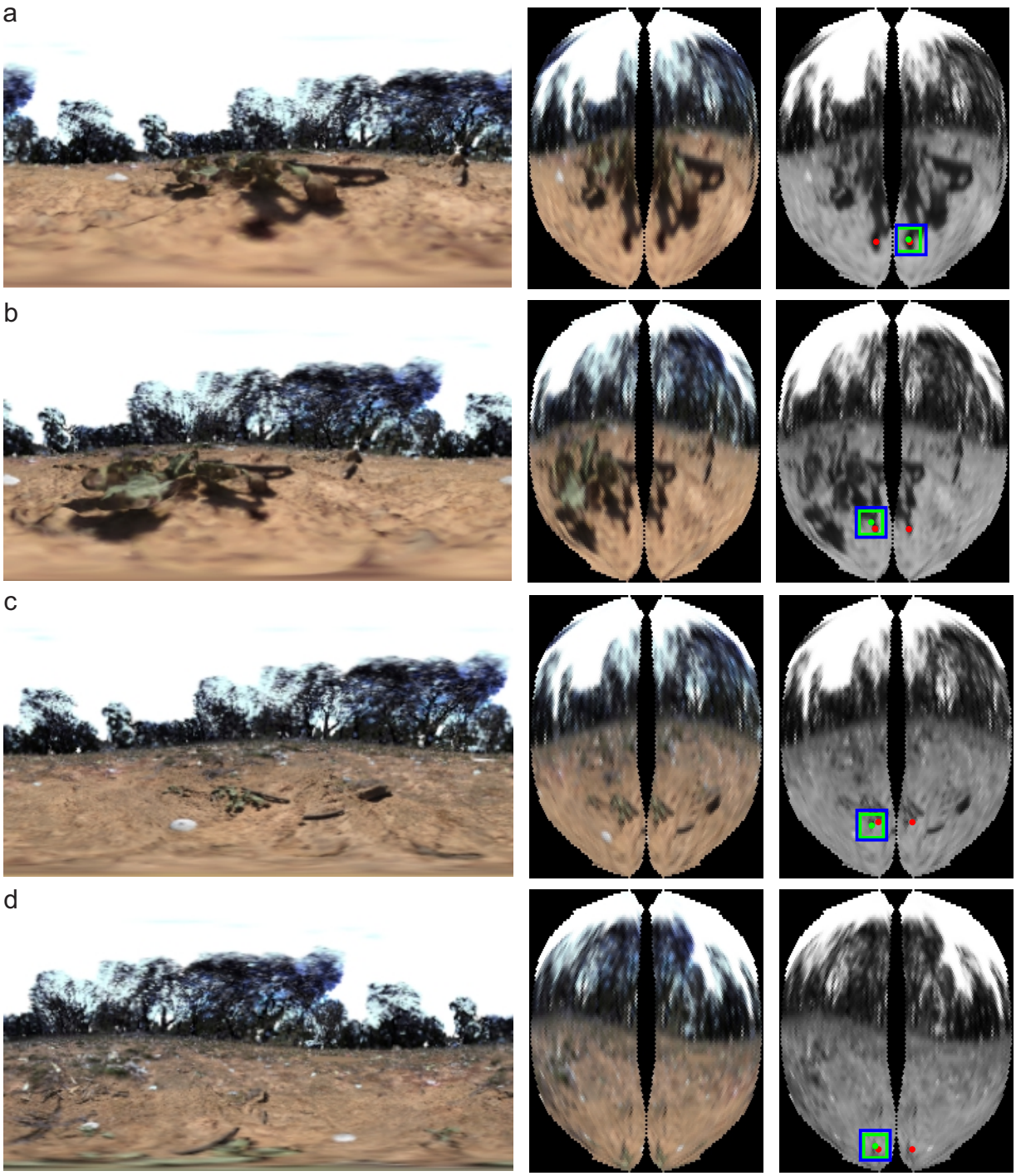


Fig.11
Stuerzl et al

Supplementary Material

Figure Legends

Figure S1 The effect of landmarks on the navigational information content of habitats. 3D models of natural environments allow us to investigate how individual landmarks, such as trees affect the topography of navigational information available in a given habitat. Top images in **a** and **b** show model views of an ant nest area before **a** and after **b** two trees closest to the nest were removed from the model. Rendered views of the scene from the nest are shown below for the two situations. **c** Maps of the transIDF for the two situations show that tree-removal makes the transIDF shallower and more symmetrical around the nest, irrespective of whether all orientations are tested (*top maps*) or the current and snapshot orientation are known (*bottom maps*).

Figure S2 How learning walks affect the range of navigational information II. Same analysis as shown in Fig. 7, but carried out under the assumption that compass information is available and that only four IDF values at each position have to be determined. At each position the current view is aligned with the four snapshots and the image difference is computed. Again, the orientation vector associated with the best matching snapshot, $\mathbf{v}(x,y) = \mathbf{v}_{i^*}$ is depicted as green arrows. The index of the best matching view is now determined from only four IDF values, $i^* = \arg\min_i \mathrm{IDF}(\mathrm{Img}(x,y), \mathrm{Img}_i)$. Blue arrows illustrate the weighted mean vectors $\mathbf{v}(x,y) = \sum_i w_i \mathbf{v}_i$, where the weights are now given by $w_i = \mathrm{IDF}(\mathrm{Img}(x,y), \mathrm{Img}_{i^*}) / \mathrm{IDF}(\mathrm{Img}(x,y), \mathrm{Img}_i)$. Otherwise conventions as in Fig. 7.

Figure S3 Camera-based reconstruction of detailed ground topography. Photograph and coloured meshes show the nesting area of ground-nesting wasps and bees. **a** Original photograph of the site with one nest location marked by an arrow. **b** Poisson mesh calculated from the input images using Structure from Motion and Poisson reconstruction. **c** Coloured and textured colour mesh generated from RGB images. **d** Textured mesh using UV grey values derived from single channel UV images. **e** False colour mesh with colour value order green-blue-UV instead of red green-blue. **f** False colour mesh with colour value order UV-green-blue instead of red-green-blue. For further details see Methods.

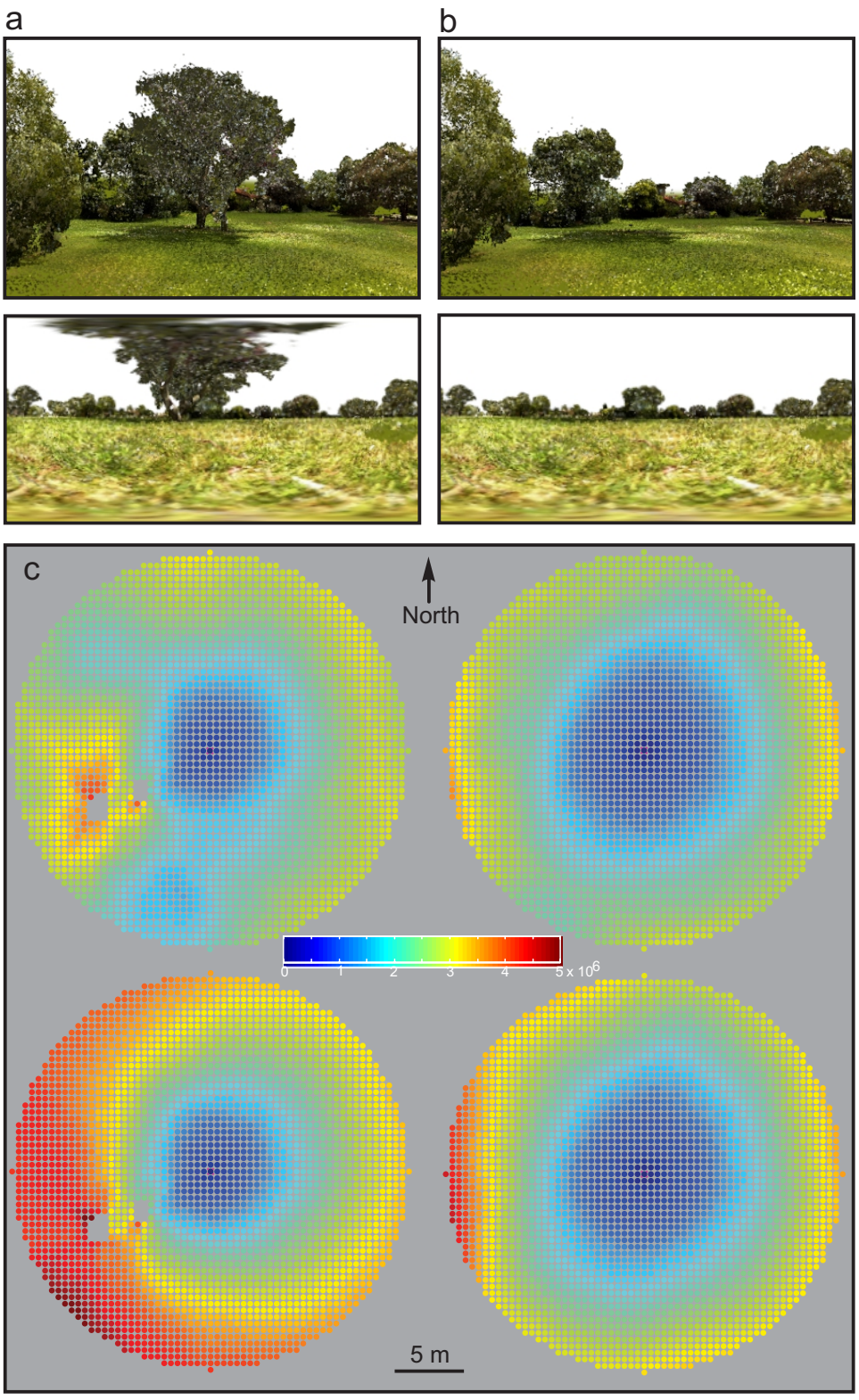
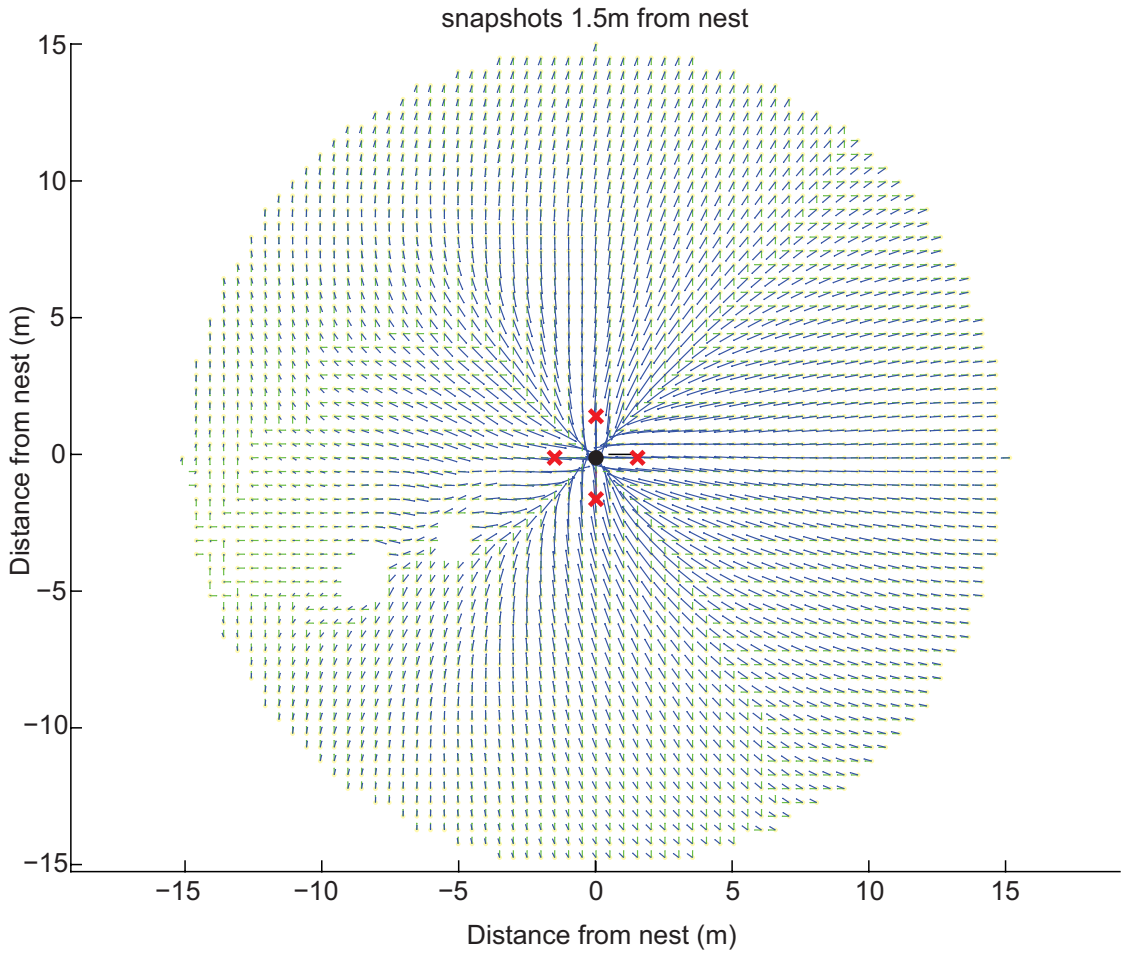


Fig. S1
Stuerzl et al



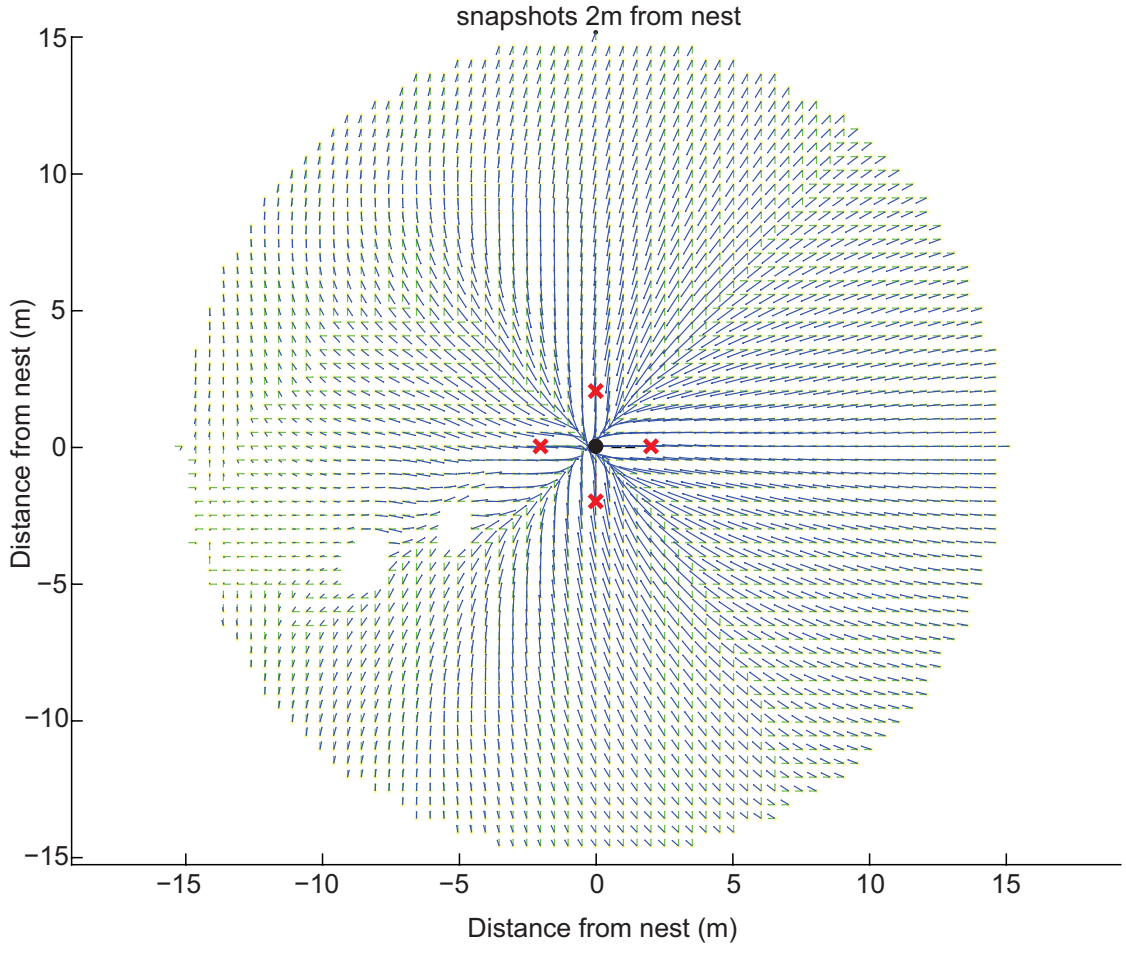


Fig. S2
Stuerzl et al

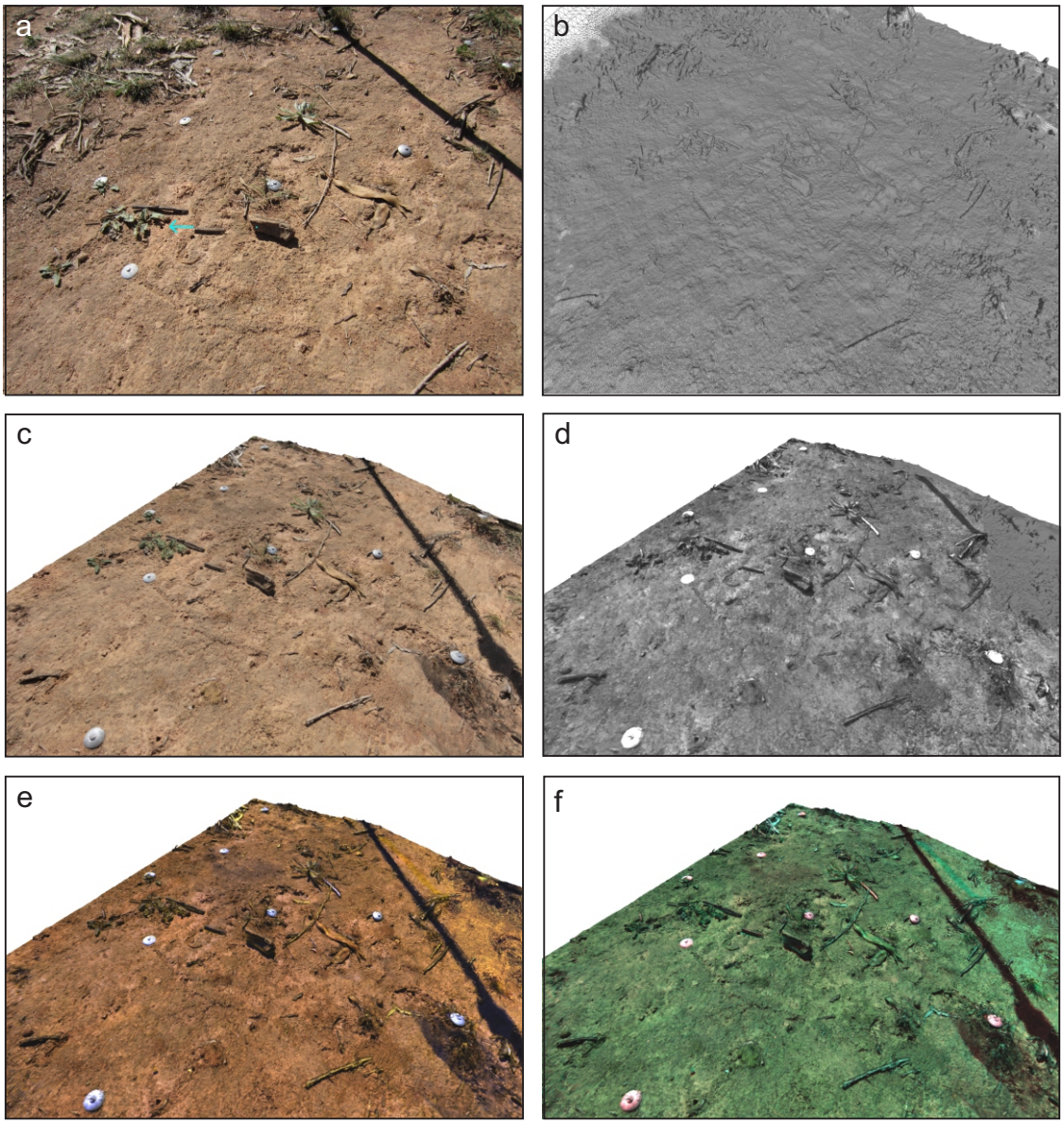


Fig. S3
Stuerzl et al